Journal of
Experimental Psychology

VOL. 28, No. 4 APRIL, 1941

CONDITIONED EXTINCTION OF A REFLEX IN THE SPINAL DOG [1, 2]

BY P. S. SHURRAGER [3] AND ELMER CULLER

University of Pennsylvania and University of Rochester

I. INTRODUCTION

In 1938 (**17**, p. 187) we presented evidence which indicated that it was possible permanently to extinguish a spinal reflex by experimental procedures other than operative. Prior to 1938, although extinction of spinal reflexes by various experimental procedures had been attempted by a number of investigators, only brief inhibition or temporary extinction had been demonstrated.[4] Sherrington (**13**, p. 218 f.) even went so far as to conclude that the nearest approach to extinction of spinal reflexes was the 'fatigue' effect observed during tetanic stimulation. However, he (as well as other investigators) was concerned with the disconcerting observation that a stimulus of weak intensity produced the fatigue effect more readily than one of high intensity.

The present paper incorporates the material of the 1938 report, together with additional data on what appears to be complete and permanent extinction of a flexion reflex of the semitendinosus muscle (S–T M.) to tail shock in the spinal dog. The criteria for 'permanent' extinction were, that the reflex (1) failed to exhibit spontaneous recovery after reasonable periods of rest; (2) did not reappear when extraneous sensory stimuli were applied [*i.e.*, was not subject to disinhibition]; and (3) did not reappear when the intensity of the stimulus which originally elicited the reflex was greatly increased.

[1] Communication No. 36 from the Laboratory of Physiological Psychology (Animal Hearing), transferred from the University of Illinois to the University of Rochester in September, 1938. This laboratory is maintained by aid of the Research Council, American Otological Society.

[2] Acknowledgment is made to Dr. H. C. Shurrager for assistance in all phases of this work.

[3] National Research Council Fellow in Psychology, University of Rochester, 1939–40. Now at the University of Pennsylvania. Publication No. 1 in a series of papers dealing with the neurological and physiological bases of conditioning completed during tenure as an NRC Fellow.

[4] As early as 1899 Partridge (**10**) had demonstrated extinction ('control') of the reflex wink.

Rather extensive data on latency of the S–T M. reflex to tail shock and to shock on the contralateral paw are also given and compared with latencies of a conditioned or learned response (*CR*) of the contralateral S–T M. to tail shock and its unconditioned response (*UCR*) to paw shock. An effort is made to show a logical relationship among the latencies of the *CR*, *UCR*, and the contralateral reflex responses to tail and paw shock, and to account for the extinction of the contralateral reflex to tail shock in terms of these latencies. An hypothesis of effector-response-latency is formulated in the light of the available categories of latencies. The exact mechanisms by which functionalization of the *CR* and defunctionalization of the contralateral reflex are accomplished, supposedly at the synapse, are still obscure, but the objective results of these mechanisms are interpreted in function terminology, so far as the experimental data permit.

II. Methodology

The operative procedure, conditioning technique, and method of recording the *CR* and reflex latencies of the spinal dog have been described in an earlier paper (Shurrager and Culler, 15, pp. 135–141). Briefly, under ether anaesthesia, the spinal cord was transected above the third lumbar spinal roots; both S–T MM. were fully exposed, free from fascia, with nerve and blood supplies intact; the insertions of the S–T MM. were cut loose and attached by threads to a pair of kymograph levers so that simultaneous records of the responses of both muscles could be made on a high-speed kymograph; the hind legs were held immobile by a drill-rod which pierced both femurs and was held rigid in clamps.

In the conditioning procedure, the conditioning stimulus (*CS*) was presented three times at one-second intervals; the third *CS* was combined with an unconditioned stimulus (*UCS*). This final combination of each set is presumably critical for the conditioning process; but the first two *CS*'s serve to show the experimenter the status of conditioning, [*i.e.*, the frequency and amplitude of the preparation's response to *CS* alone]. The *CS* was a minute shock to the tail, while the *UCS* was a shock to the hind paw strong enough to cause maximal contraction of the S–T M. The *UCS*, in the experiments here reported, was never omitted, even when the animal's *CR* to the *CS* had reached 100%.[5]

The latencies of the responses of left and right S–T MM. to *CS* and *UCS* were calculated from the kymograph records. The significance of the differences among them was in each case determined by Fisher's method for small numbers (7, chap. 5). All estimates of '*P*' are taken from 'Student's' table (19, 105 ff.), with due regard for degrees of freedom.

The loci of the *CS*-reflex and the *CR* elements in the two S–T MM. were carefully observed.

III. Results

In some of the spinal preparations studied, brief stimulation of the tail with 10–30 V. of AC shock normally elicited a persistent [*i.e.*, continuing through as many as 200 preliminary test trials] reflex contraction of one S–T M. while the other S–T M. remained flaccid. This inherent reflex response to the *CS* was in the majority of cases located in the right muscle. When a reflex response to tail-stimulation was inherently present in one S–T M., a *CR* of the other

[5] This conditioning procedure [*i.e.*, the *UCS* is applied in the conditioning sequence regardless of the preparation's reaction or failure to react] is of the 'classical' form (8, 27).

muscle to the same stimulus could be established by the process of conditioning. If the conditioning procedure was continued for a great number of trials, the *CR*, as it became more firmly established, tended to inhibit the contralateral reflex, so that eventually the latter completely disappeared.

The complete and permanent failure of the contralateral reflex as a result of pitting against it a *CR* of the corresponding muscle to the same (tail) stimulus will henceforth be referred to as *conditioned extinction* of the reflex. The course of such conditioned extinction was carefully observed in five spinal dogs and the record of one of them (which appeared to be representative of all five) is given with some detail in Fig. 1. Here the reflex response is shown to be at

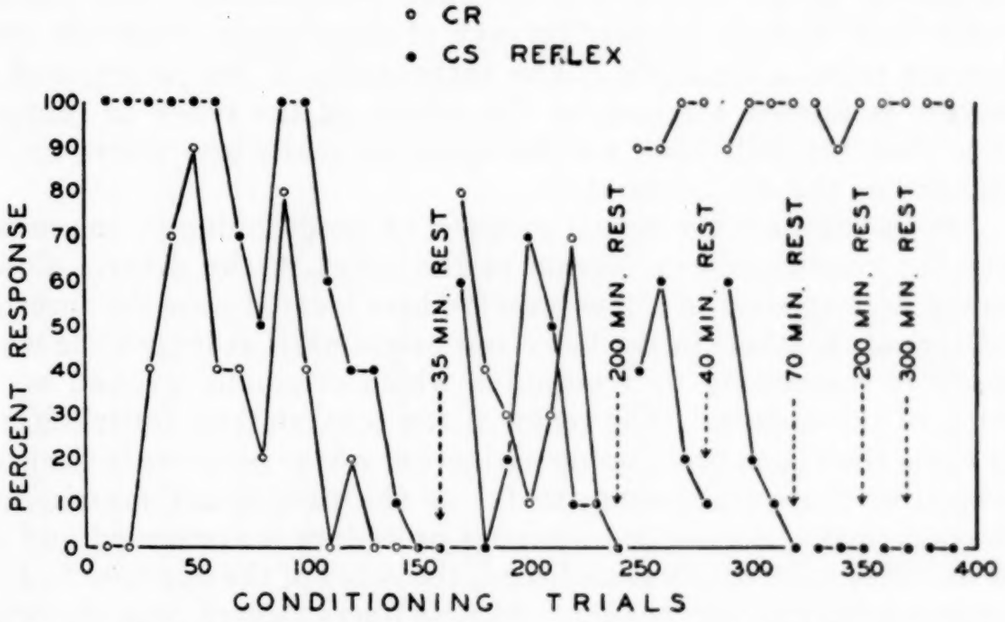

FIG. 1. Record of conditioned response, *CR*, to tail-stimulus in the left M. Semitendinosus in one spinal dog and simultaneous conditioned extinction of an innate reflex, *CS*-reflex, in the right M. Semitendinosus to the same tail stimulus. Both conditioned response (open circles) and conditioned extinction (solid circles) are expressed in percent response in each successive ten trials.

100 percent in the first twenty trials, while the *CR* is completely lacking. At 50 trials the *CR* was present 90 percent of the time (nine trials out of ten), while the reflex was still at 100 percent. Between 60 and 150 trials the two responses seemed to become highly instable and mutally inhibitive, so that at 150 trials both had disappeared. Here a rest period of thirty-five minutes was interposed. By the end of this period partial recovery occurred in both reflex and *CR;* they were now responding at 60 percent and 80 percent respectively. For the next seventy trials the reflex and *CR* again became instable and mutally opposed [*i.e.*, when one responded, the other failed] and

at 250 trials both were once more completely absent. After a rest period of two hundred minutes the *CR* responded eighteen times out of the first twenty trials, whereas recovery of the reflex was much less complete. At 280 trials the *CR* had been present for twenty straight trials, while the reflex had sunk, to 10 percent response. After a rest period of forty minutes the *CR* to tail stimulus failed once, and the reflex four times, in the first ten trials. Between 290 and 320 trials the *CR* failed but once, while the reflex dropped to zero. At 320 trials the reflex seemed to have been finally extinguished, for interpolated rest periods of seventy, two hundred, and three hundred minutes effected no recovery in it.

The conditioning procedure was continued through 390 trials with the animal whose record is graphically represented. The *CR* was absent only once in the last seventy of these trials, while the reflex response failed completely. The authenticity of the conditioned extinction is further attested by the failure of the reflex to reappear even when the tail shock was increased to nearly 400 volts (the full intensity of the AC stimulator).

Let us compare the actual progress of conditioning in one muscle with the conditioned extinction of the reflex in the other. Conditioning first appears in a few muscle-fibers located near the insertion and spreads to fibers in the belly and origin until at length the whole muscle is responding to a stimulus which originally evoked no reaction of the muscle. The reflex of the contralateral muscle has at the same time been undergoing extinction whose progress is the exact reversal of the conditioning, so far as the participant muscle-units are concerned. As the conditioning procedure is continued and the *CR* becomes more firmly established, the reflex of the opposite S–T M. becomes progressively weaker. Muscle fibers located near the origin are first to fail, fibers located in the belly of the muscle next, and those at the insertion last. It should be noted that this course of conditioned extinction parallels the course of extinction in a spinal *CR*.

There seems to be a marked gradation from insertion to origin in the ease with which different parts of the muscle are conditioned and in the stability of the *CR*, once it is established. Since that part of the muscle near the insertion is first to condition and last to extinguish, it may be that, of the motoneurone population, those innervating this group of muscle-fibers form synaptic associations more readily and maintain them better than those innervating other parts of the muscle. Extinction, then, would affect the muscle-fiber groups inversely to their degree of synaptic stability within the cord (Shurrager and Shurrager, 18).

A comparison of latency-values of the reflex response of one S–T M. with the latencies of the contralateral muscle's *CR* to tail-

shock and its *UCR* to paw-shock reveals some interesting differences among them. Latency-measures for seventeen cases of paired flexions in one dog are given in Table I. In those seventeen cases both the *CR* and the *CS*-reflex occurred upon both presentations of the *CS* in the conditioning procedure. When either the reflex or the *CR* failed to respond to one or both *CS*'s in a given trial, that case was not included in the table; since the state of the organism during a trial in which there appeared a response to each *CS* by each muscle [*i.e.*, four recorded responses] did not seem strictly comparable to its state during a trial in which only three or two recorded responses were obtained. The figures given in Table I for *CR* and *CS*-reflex

TABLE I

LATENCY IN MILLISECONDS OF THE CONDITIONED RESPONSE (*CR*), OF THE CONTRALATERAL REFLEX TO *CS* (*CS-R*), AND OF THE UNCONDITIONED REFLEX (*UCR*) FOR SEVENTEEN TRIALS IN A SINGLE SPINAL DOG

Case No.	*CR*	*CS*-Reflex	*UCR*	*CR–CS-R*	*CR–UCR*	*CS-R–UCR*
1	37	43	15	−6	22	28
2	36	45	23	−9	13	22
3	46	37	17	9	29	20
4	43	49	26	−6	17	23
5	46	45	20	1	26	25
6	49	55	18	−6	31	37
7	25	34	18	−9	7	16
8	31	37	34	−6	−3	3
9	50	46	20	4	30	26
10	23	32	29	−9	−6	3
11	49	56	21	−7	28	35
12	46	64	21	−18	25	43
13	42	42	24	0	18	18
14	34	49	25	−15	9	24
15	45	60	26	−15	19	34
16	37	55	20	−18	17	35
17	37	55	19	−18	18	36
Total.....	676	804	376	−128	300	428
Mean.....	39.8	47.3	22.1	−7.5	17.7	25.2
Median.....	42.0	46.0	21.0	−9.0	18.0	25.0
S.D.$_{dis.}$.....	8.22	9.29	4.79	7.88	10.96	11.25
S.E.$_{Mean}$.....	1.99	2.25	1.23	1.91	2.56	2.72
t.....				3.93	6.68	9.33
P.....				.01>*P*>.001	<.001	<.001

are in each case the means of the latencies for the responses to first and second *CS*. In this animal the reflex was inherently present in the right S–T M. and a *CR* was established in the left S–T M. Columns *2* and *3* in Table I exhibit respectively the latencies of responses from the left and right S–T MM. when activated by a single shock to tip of the tail. In column *4* is the latency of the response of the left S–T M. (the conditioned muscle) to the *UCS* (shock to paw).

It is apparent in Table I that the *CR* in the trained muscle has a consistently shorter latent period than does the reflex-response of the

corresponding contralateral muscle, although both are elicited by the same stimulus; and that the latency of the *UCR* of the trained muscle is shorter than either *CR* or *CS*-reflex. That the differences between the *CR* and the *CS*-reflex of the contralateral muscle, between *CR* and *UCR* of the conditioned muscle, and between the *UCR* of the conditioned muscle and the contralateral *CS*-reflex are all reliable was established by solving for '*t*.' In all cases, $P < .01$.

The latencies given in Table I for *CR* and *UCR* of the conditioned muscle and for reflex response of the contralateral muscle are graphically represented in Fig. 2. The shaded areas indicate the actual

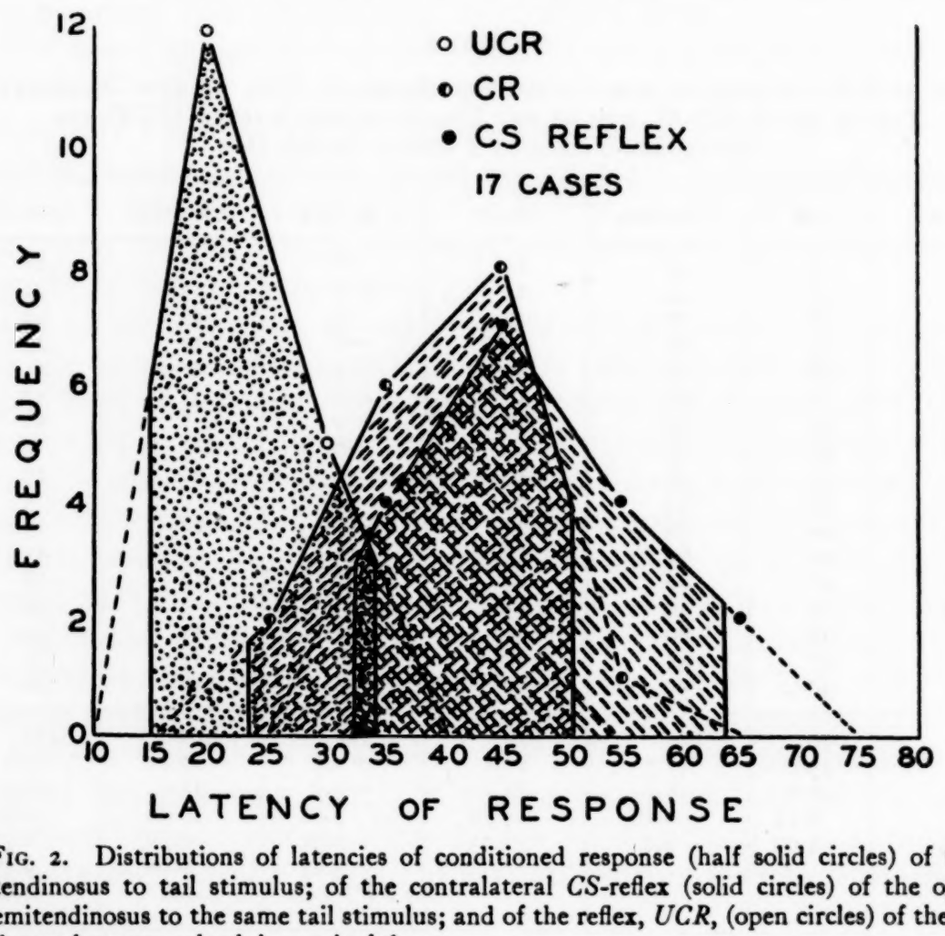

FIG. 2. Distributions of latencies of conditioned response (half solid circles) of the M. Semitendinosus to tail stimulus; of the contralateral *CS*-reflex (solid circles) of the opposite M. Semitendinosus to the same tail stimulus; and of the reflex, *UCR*, (open circles) of the conditioned muscle to paw shock in a spinal dog.

distribution of cases for each type of response. Dotted lines indicate the extreme range of cases according to class intervals. That the *CR* latency lies somewhere between the latencies of *UCR* and *CS*-reflex is here quite obvious.

We are unable to present a sampling of latency data from more than one of the five dogs in which conditioned extinction was carried to completion, since in only one of these animals were the responses recorded by high-speed kymograph. In order to determine whether

the latency measures of *CR* and *UCR* obtained from this one animal may be accepted as representative, *CR* and *UCR* latency values were calculated from the records of a number of other spinal preparations. These values are given in Table II, which is made up of latency data based on one hundred and forty-nine readings from twelve spinal dogs.

In Table II the latency-figure for a given *CR* is actually the average of the muscle-response latencies to the first two *CS*'s. The la-

TABLE II

COMPARATIVE LATENCY IN MILLISECONDS OF THE CONDITIONED RESPONSE (*CR*) AND OF THE UNCONDITIONED REFLEX (*UCR*) FOR THE FLEXION OF M. SEMITENDINOSUS IN TWELVE SPINAL DOGS

Dog No.	No. Cases Flexion	*CR* Mean	*UCR* Mean	*CR*–*UCR* Mean
1	4	37.3	25.5	12.8
2	9	99.2	28.7	70.5
3	5	133.4	27.2	106.2
4	13	117.8	42.4	75.4
5	4	56.8	30.5	26.3
6	29	68.2	30.5	37.7
7	22	71.5	30.3	41.2
8	2	98.5	29.0	69.5
9	4	56.5	50.5	26.0
10	3	56.7	29.7	27.0
11	24	69.8	31.0	38.8
12	30	43.1	22.5	20.6
Total	149	908.8	356.8	552.0
Mean Dog		75.7	29.7	46.0
Mean Case		71.0	29.6	41.4
Median Dog		69.0	30.0	38.2
Median Case		61.0	29.0	32.0
$S.E._{Mean}$		8.66	3.03	8.12
$S.D._{dis.}$		30.0	10.5	28.11
t				5.66
P				<.001

tency of the *UCR* is that of the response of the trained muscle to the *UCS* (paw shock), which follows the second *CS* after a one-second interval and is combined with a third *CS*. Column *2* of Table II gives the number of recorded flexions from each dog. Column *3* gives the average latency of the *CR*, column *4* the average latency of the *UCR*, and column *5* the difference between the averages. The means and medians of the distributions show that the latent period of the *CR* to the *CS* is markedly longer than that of the *UCR* to the *UCS*. Fisher's method for determining the reliability of a difference yields a '*t*' value of 5.66, which means that there is less than one chance in a thousand that this difference would occur in a random population ($P<.001$). The standard deviations of the means of the *CR* and *UCR* latency distributions reveal a much greater interna consistency in the distribution of *UCR*'s.

The data given in Table II are shown graphically in Fig. 3. The distributions of *CR* and *UCR* latencies are plotted in eight class intervals, using the same base line for both distributions. The class interval for the *UCR* latency-distribution is eight msec. while the

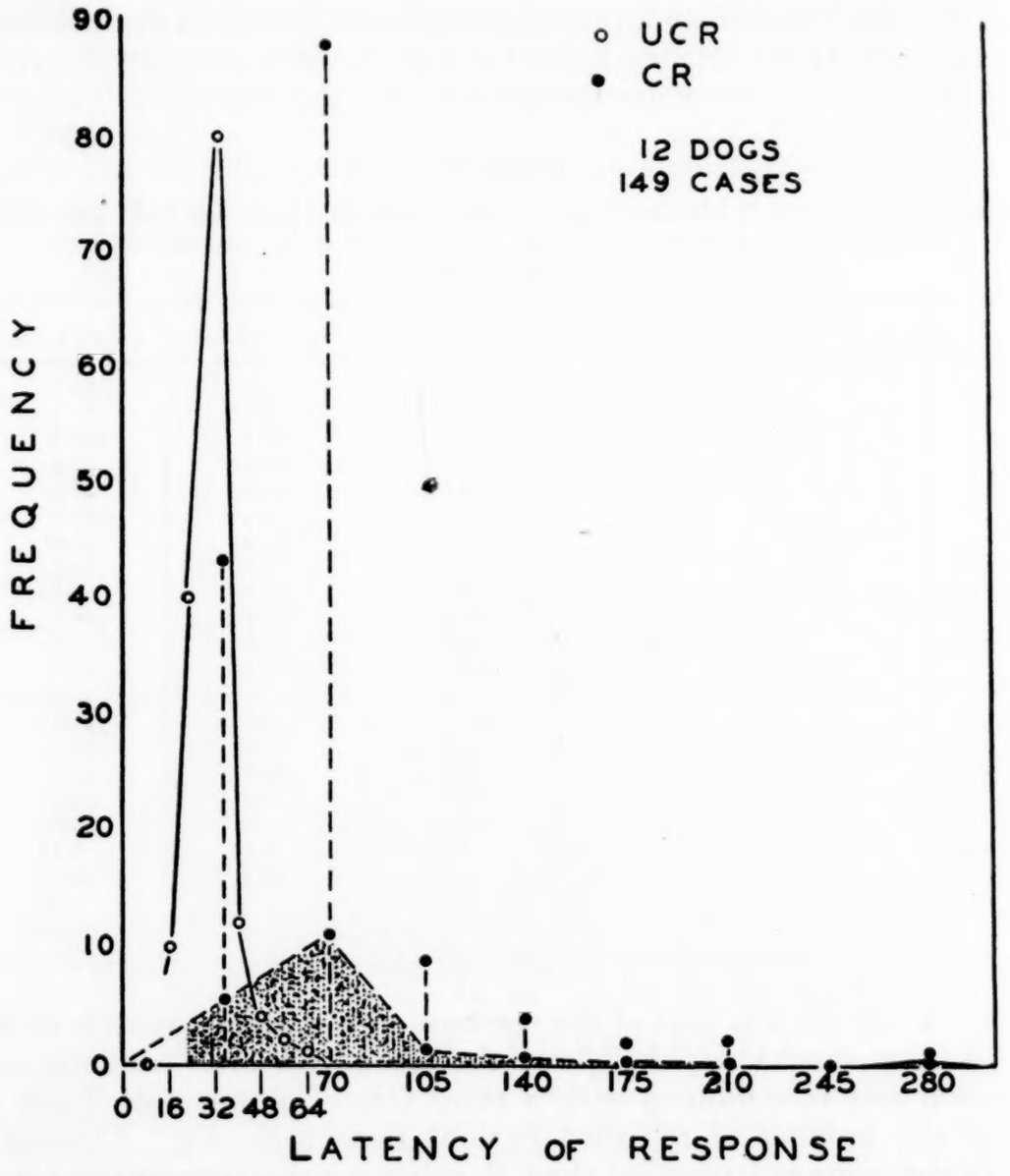

FIG. 3. Distributions of latencies in milliseconds of the unconditioned reflex, *UCR*, (open circles) and of the conditioned response, *CR* (solid circles), from the flexion of M. Semitendinosus in twelve spinal dogs.

class interval for the *CR* distribution is thirty-five msec. This difference in size of class interval makes the area under the curve representing the distribution of *CR* latencies proportionally much greater than the area under the *UCR* curve. In order to compare the two distributions more readily, the area under the *CR* curve was reduced until it equalled the area under the *UCR* curve. This reduced *CR* area is represented by the shaded portion of the diagram. The lines

(solid for *UCR* and broken for *CR*) enclosing the distributions are stopped at the abscissal values at which the cases cease to appear.

The means for the twelve dogs of Table II are in general larger than those of Table I but otherwise similar. In both instances the *UCR* latency is markedly more consistent and reliably shorter than the *CR* latency.

TABLE III

LATENCY IN MILLISECONDS OF THE HOMOLATERAL (HL) AND CONTRALATERAL (CL) FLEXION REFLEXES ELICITED BY A SINGLE SHOCK TO ONE HINDPAW IN FIFTEEN SPINAL DOGS

Animal No.	No. Cases Flexion	Homolateral Mean	Contralateral Mean	*CL–HL* Mean
1	33	18.8	179.2	160.4
2	7	16.0	266.3	250.3
3	18	25.6	158.7	133.1
4	6	26.0	116.2	90.2
5	13	29.7	109.5	79.8
6	4	20.3	25.3	5.0
7	6	24.0	33.5	9.5
8	7	27.6	76.1	48.5
9	4	23.5	154.0	130.5
10	19	29.9	111.5	81.6
11	10	22.1	30.2	8.1
12	6	27.0	81.7	54.7
13	21	31.2	118.1	86.9
14	11	30.3	107.5	77.2
15	5	39.2	54.7	15.5
Total	170	391.2	1622.5	1231.3
Mean Dog		26.1	108.2	82.1
Mean Case		25.9	123.6	97.7
Median Dog		26.0	109.5	83.5
Median Case		25.0	115.0	90.0
$S.D._{dis.}$		5.74	61.42	66.76
$S.E._{Mean}$		1.70	15.86	17.25
t				4.76
P				<.001

A final comparison of latency-values can be made between those of the flexion-response of the contralateral muscle to the *UCS* and the flexion-response of the homolateral muscle to the same stimulus. The fact that there is a response by the contralateral flexor muscles to stimulation of one paw has received little attention in previous studies of reflex activity. This is not strange, since, when the leg is intact, its extensor-response to contralateral stimulation is so strong as to obscure the less obvious contraction of the flexors. However, the very nature of the physiological mechanism presupposes a certain amount of flexion in opposition to any extension, if control of the muscle groups involved is to be at all accurate. When the insertion of the S–T M. contralateral to the *UCS* is cut free and attached to a kymograph lever, the normally obscured flexion of the muscle can be recorded. If, as in the present study, simultaneous records are taken from both muscles, comparison of latencies of homo-

lateral and contralateral flexion responses to the same stimulus can be made. The contralateral flexion is, of course, diminutive in comparison with homolateral flexion.

The reflex response of the spinal animal (or of an unexcitable normal animal) to the *UCS* to the paw can be expressed thus: homolateral flexion > contralateral flexion, contralateral extension > homolateral extension. Presumably these spinal-reflex responses are organized into the final *CR* when *UCS* is combined with *CS* in the standard conditioning procedure. The *CR*, then, consists of a balanced flexion of the conditioned muscle (or muscle group) in which flexion > extension and a concurrent balanced contralateral extension in which extension > flexion.

Table III presents the latency values of 170 paired flexions to the *UCS* in fifteen spinal dogs. The mean homolateral latency is 25.9 msec. and the mean contralateral latency is 123.6 msec. There is less than one possibility in a thousand that this is a 'chance' difference. That the shorter homolateral latency is much more consistent than the contralateral latency is indicated by the striking difference between the standard errors of the means of the two distributions.

The distributions of latencies of the homolateral and contralateral reflexes simultaneously elicited by the *UCS* are shown by the frequency polygons of Fig. 4. Solid lines indicate the distribution of

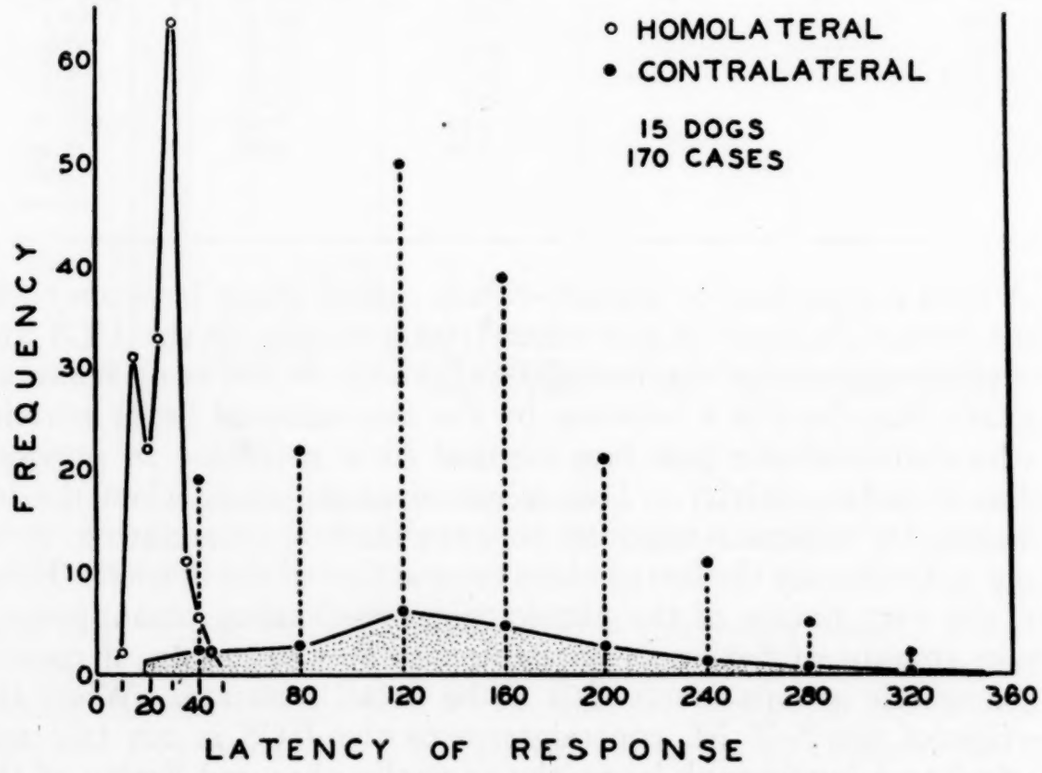

FIG. 4. Distributions of latencies in milliseconds of the homolateral (open circles) and contralateral (solid circles) flexion of MM. Semitendinosus elicited by a single shock to one hindpaw in fifteen spinal dogs.

latencies for the 170 homolateral flexions and broken lines for the contralateral flexions.

IV. *Discussion*

The closest approximation to true extinction of spinal reflexes hitherto reported [*i.e.*, prior to 1938 (**16**)] occurs in a series of experiments reported by Prosser and Hunter (**11**, 609, ff.) in 1936. They attempted to extinguish the reflex startle response of the normal rat and certain spinal reflexes in the spinal rat by repeated application, at short intervals, of the stimuli eliciting these reflexes. With normal animals the stimulus was an auditory click to which the rat replied with a reflex startle (leg-response). The spinal reflexes studied were (1) leg-response to tap on rump or tail; (2) flexion of tail when tapped near its tip; (3) tail- and leg-response upon tetanic stimulation of tail; and (4) crossed leg-response to direct stimulation of the contralateral saphenous nerve. The extent of the response was recorded by means of a concentric needle electrode placed either in the ventral base of the tail or in the gastrocnemius muscle of one hind leg. Both the startle response of the normal rat and the spinal reflexes designated above were observed to weaken and disappear after repeated applications of their respective stimuli at intervals of ten to fifteen seconds. Extinction of these reflexes consisted of a gradual diminution of the number of motor units involved in the response and a decrease in the duration of the after-discharge.

Extinction was, however, temporary in both normal and spinal rats. In each case, spontaneous recovery was found to occur immediately after extraneous stimulation ('disinhibition'), after interpolated rest periods ranging from five to thirty minutes, and, in the case of spinal animals, when the stimulus-intensity was increased. No permanent extinction in either type of preparation was reported.

The course of conditioned extinction of the S–T M. reflex to tail stimulation in the spinal dog exhibits certain resemblances to the temporary extinction observed by Prosser and Hunter in rats. The representative case portrayed in Fig. 1 shows that in spinal dogs the reflex exhibited a partial degree of spontaneous recovery when rest periods were interpolated early in the extinction process, whereas in spinal rats recovery was complete and consistent. A further similarity appears in the manner in which the reflex gradually extinguished and eventually failed. In both cases the muscle response diminished motor-unit by motor-unit until it completely disappeared.[6] Despite

[6] In the exposed S–T M. of the spinal dog, unit-by-unit decrease in the response of the whole effector population can be observed. The records of reflex-extinction in the rats were based upon the responses of only those few fibers affecting the tip of the concentric needle electrode. Presumably, if the electrode was located near the origin of the responding muscle, the number of

these resemblances, the two processes differ notably in the permanence and stability of extinction. It seems that the mechanism responsible for this complete extinction of the S–T M. reflex in the spinal dog somehow involves the contralateral *CR* and depends in some degree upon the procedure used in extinction. In other words, the 'permanent' type of extinction here reported seems to depend upon an inhibitory interaction between bilateral responses or their mediating mechanisms. Several clues to the physiological basis of this inhibitory interaction may be mentioned.

Clue I: Evidence for bilateral motor inhibition existing in isolated spinal preparations was presented (Shurrager and Culler, **17**, p. 126) in 1939. This paper was concerned with the effect upon spinal reflexes of severing the contralateral motor roots and showed that section of these motor roots released complex repetitive reflexes of the homolateral (stimulated) side in response to brief single shocks to one paw. Prior to severing these contralateral motor roots, the same stimulus elicited only a single normal flexion and subsequent relaxation of the stimulated leg. Since the afferent roots had been previously severed, it was suggested that the motor roots, in addition to their efferent function, mediated some afferent inhibitory effect upon the contralateral reflexes [7] (Windle, **20**, and Schäfer, **12**). In any event, the inhibitory effect described was in some way dependent upon presence of the contralateral motor roots; but their exact role in the mechanisms of conditioned extinction is not at present known.

Clue II: If some unit-for-unit opposition exists between the motor elements participating in the S–T M. reflex to tail shock and those participating in the contralateral *CR*, it must be an opposition of dissimilar (anatomically unsymmetrical) units. Such unit-for-unit opposition would imply, for example, a mutually inhibitory interaction between the insertion units of one muscle and the origin units of the other, since the muscle fibers near the insertion are first to condition and those at the origin are first to extinguish.

A more plausible assumption is that the mutually inhibitory effect between the two responses (*CR* and reflex) results from a mass

---

trials required for apparent extinction would be smaller than if the electrode were placed among fibers in the belly or insertion of the muscle, since evidence from the dogs indicates that extinction occurs first among fibers near the origin and proceeds toward the insertion.

[7] Such an afferent inhibitory effect could be mediated by: (1) the afferent fibers suggested by various investigators to occur in the ventral roots; (2) some bioelectric gradient between muscles and cord, which might operate to maintain a balanced inhibitory relation between the two halves of the cord. This balance would presumably be destroyed by severing the motor roots. That the conditioning and extinction curves in Fig. 1 are not the typical curves of spinal conditioning (**14**; **15**, p. 149) is probably due to the interaction of the opposed processes (conditioning and conditioned extinction) rather than to any fundamental differences among the mechanisms involved. It should be noted that none of the animals included in the original spinal conditioning curves exhibited the contralateral reflex response to tail-shock.

interaction (pool effect) of the bilateral motor-unit populations which participate in the responses. If either assumption be true, the inhibitory effect cumulates with each conditioned extinction trial until it eventually causes overt failure of the reflex.[8]

Clue III: That the respective *latencies* of the responses in bilaterally interactive systems may be significant in the extinction of spinal reflexes is suggested by the fact that, although *UCS* is applied only on one side, it elicits flexion of *both* S–T muscles. We might, therefore, expect the paw-shock not only to operate in establishing the homolateral *CR* but also to modify the *CS*-reflex of the contralateral muscle, which is already present. Consider again the Prosser-Hunter temporary extinction of reflexes in rats. Their stimuli produced bilateral responses (startle) in both normal and spinal animals. Perhaps the bilaterality of the responses, which were recorded from one side only, was an operative factor in bringing about extinction. That is, one side may have responded prior to the other and thus operated in a manner comparable to the conditioned extinction in spinal dogs. In the rats, perhaps because no unilateral incentive (*e.g.*, paw shock) was applied forcibly and consistently, neither side achieved enough functional superiority to effect permanent extinction.

Drabovitch's summary (**6**, 93 ff.) of several studies (Chauchard, Chauchard, and Drabovitch, **3**, **4**, **5**) on chronaxie as related to conditioning suggests one interpretation based on latencies. He and his co-workers found that the chronaxies of both flexor and extensor muscles increase to as much as twice the normal resting values during the period of three or four seconds from onset of the *CS* (bell) until conditioned leg flexion occurs and again for a short time thereafter; but that they decrease below the normal resting value during actual execution of the *CR*. Parallel changes in chronaxie, before, during and after the delayed *CR*, occur at the cortical motor point for the conditioned limb. In spinal dogs the bilateral flexions, with short latency on same side as paw-shock and longer latency on opposite side, afford the needed prerequisites for bilateral conditioning. With each muscle, a *CR*, therefore, tends to arise just prior to the corresponding *UCR*.[9] But in the contralateral muscle (longer latency) the picture is complicated by the *CS*-reflex which occurs during the latent period of the presumptive *CR*. According to Drabovitch, the chronaxie of the reflex-mechanism rises during the latent period pre-

[8] The progress of conditioning the *CR* and extinguishing the *CS*-reflex and the stability of the conditioned and extinguished states, once they are established, respectively resemble the *central conditioning* and *central extinction* and the *conditioned* and *extinguished steady states* described in earlier papers (**14**; **15**, 151ff.).

[9] It has long been known that a *CR* tends to be delayed until or near the onset of *UCS*. In a special case like the present, it seems reasonably to assume that *CR* would likewise tend to occur at or near the time of *UCR* in the same muscle.

ceding an established *CR*; from which we may infer that the supranormal chronaxie would diminish the excitability of the *CS*-reflex mechanism and lead to its eventual extinction.

A second interpretation may be based on the latency of the effector-response (*UCR*). In the final step of each conditioning trial, *CS* and *UCS* are presented together, one to tail and one to a hindpaw. The mean latency of the *CS*-reflex is 40 to 50 msec. (Table I) and of the *UCR* (crossed flexion reflex) in the same muscle, 110 to 120 msec. (Table III, column *4*). In the present case, the figures of Table I, column *3*, intimate (the available evidence is far too limited to prove) that the latency of *CS*-reflex begins to rise in the later trials. If this be a true case of conditioned delay (the original *CS*-flexion being conditioned in latency toward that of the crossed *UCR*-flexion), it suggests that the *CS*-reflex may extinguish because it is temporally displaced;[10] that is, with increasing latency it loses stability and amplitude, and eventually vanishes. Similar relations were found by Kellogg and Walker (**9**) with conditioned flexion in dogs, by Campbell and Hilgard (**2**) with conditioned eyelid responses in man, and by Campbell (**1**) with the conditioned knee-jerk in man (cf. Hilgard and Marquis, **8**, p. 138).[11]

In brief, the question is: in this series of *CS–UCS* combinations, what differentiating mechanism operates to bring about conditioning one one side of the animal and extinction on the other? On the sensory side there seems to be no clue to a differential effect. Since the *CS* is applied to the tip of tail, the temporal sensory aspects of the *CS* should be virtually identical on both sides of the cord. Although *UCS* is unilateral in that there are no primary contralateral sensory elements, it is a common stimulus to both homolateral and contralateral *UCR*'s. On the motor side, however, we have shown that the flexion response of the homolateral muscle to the *UCS* is markedly shorter than that of the contralateral muscle. There is a difference of 90 to 100 msec. between the mean latencies of the two flexion responses (Table III). This difference between the homolateral and contralateral latencies must presumably originate as differential delays within the cord, since peripheral nerve transmissions are relatively fast and constant. The difference in *effector-response-latency* of the two reflex responses to the *UCS* may account for the divergent effects of the *CS–UCS* presentations in the two S–T

[10] Recall the Prosser and Hunter found no latency change for the individual units during their extinction and therefore this explanation does not seem applicable to their temporary extinction.

[11] Sherrington's comment, "On the whole, the 'long' spinal reflexes are more variable and less validly predictable than the short" (**13**, p. 166) suggests that true spinal reflexes, as well as conditioned responses, also become less stable when the response is delayed. As their latency rises, they are more likely to fail.

muscles.[12] The latency of the *UCR* on the homolateral side is comparatively very short. Since *CS* and *UCS* are presented together this short *UCR* latency should operate to produce a short latency of the *CR* and thereby establish it firmly. The latency of the contralateral response to the *UCS* is quite long. This might well operate to condition a delay in the contralateral reflex response to the *CS* until just prior to [*i.e.*, as nearly coincident with it as the conditioning procedure permits] the *UCR*. This conditioned delay might, in turn, contribute to the inhibition and final extinction of the reflex, since delayed responses tend to extinguish rapidly.[13] In general the functional relation between *CS* and *CR* may be determined by the unequal response-latencies of the various muscles affected by the *UCS*.

V. *Summary*

1. Complete and permanent *conditioned extinction* of a spinal reflex is demonstrated in the spinal dog. In some spinal dogs a reflex response of the flexor muscle, M. Semitendinosus (S–T M.), was found to be inherently present on one side and absent on the other. When electric or mechanic stimulation to the tail (*CS*) was combined with shock to the hindpaw on the side exhibiting no reflex to tail shock, a conditioned response (*CR*) of the S–T M. on that side could be established. As the conditioning procedure was continued and the *CR* became more firmly established, the *CS*-reflex of the contralateral muscle was observed to weaken and disappear. The reflex did not reappear when the *CS*-intensity was greatly increased, when extraneous stimuli were applied, nor when extensive rest periods were interpolated in the conditioning process.

2. The hypothesis is advanced that the permanent extinction of a spinal reflex here reported is somehow dependent upon an inhibitory interaction between the contralateral *CR* and *CS*-reflex mechanisms.

[12] This theory of the role of *effector-response-latency* may account for the progression of the *CR* from a generalized to a specific response during the process of conditioning the normal animal. In the early stages of conditioning, the *UCR* begins within the member, *e.g.*, leg, directly affected by *UCS* and spreads to other muscles (general activity). The *CR* at first also involves this general response of the animal. As conditioning is continued and the *CR* becomes more specific, those elements of the general response which are farthest from the *CS* in response-latency tend to drop out first. Eventually the *CR* comes to include only those elements of the *UCR* which were temporally most closely associated with the *UCS*, and consequently with the *CS*. The same latency factors operate during extinction of the *CR*. The response to the *CS* tends to become more and more circumscribed until, in the final stages, only a flick of the immediate member originally subjected to the *UCS* can be observed.

[13] This phenomenon of delay is the first sign of extinction of the *CR* in normal animals and regularly precedes the eventual failure. The progressive delay in extinction, however, is not observed when the animal has been repeatedly reëxtinguished. It may reach a state in which reëxtinction may be effected in a single trial. Under these conditions no delay is observable. Such a phenomenon was observed in spinal reconditioning and reëxtinction in the spinal dog (14; 15, p. 151).

3. The latencies of the respective responses of the two S–T MM. to *CS* and *UCS* arranged in order from shortest to longest are: (1) *UCR* of homolateral muscle; (2) *CR* of homolateral muscle; (3) *CS*-reflex of contralateral muscle; (4) *UCS*-reflex of contralateral muscle. It is suggested that the response system which has a *shorter latency* enjoys a functional superiority in the bilateral interactions which may underlie conditioned extinction.

4. Both the conditioned extinction of the *CS*-reflex and the establishment of the *CR* are presumably effected at the time when *CS* and *UCS* are presented together. It is suggested that the differentiating mechanism which operates to bring about conditioning on one side of the animal and conditioned extinction on the other is mediated by the motor elements, since the temporal sensory aspects of *CS* and *UCS* are presumably identical on both sides. The latency of the homolateral being markedly shorter than that of the contralateral flexion, this difference in *effector-response-latencies* of the two reflex responses to *UCS* may account for the different effects of the same *CS-UCS* combinations upon the two S–T muscles.

(Manuscript received October 7, 1940)

References

1. Campbell, A. A., The interrelations of two measures of conditioning in man, *J. exp. Psychol.*, 1938, **22**, 225–243.
2. ——, & Hilgard, E. R., Individual differences in ease of conditioning, *J. exp. Psychol.*, 1936, **19**, 561–571.
3. Chauchard, A. & B., & Drabovitch, W., Réflexes conditionnés et chronaxie, *C. R. Soc. Biol.*, Paris, 1934, **116**, 959–962.
4. ——, & ——, Modifications de l'excitabilité des neurones périphériques au cours du mouvement, dans le réflexe conditionné, *C. R. Soc. Biol.*, Paris, 1935, **119**, 76–78.
5. ——, & ——, La chronaxie de l'écorce cérébral aux divers temps du réflexe conditionné. Isochronisme des neurones centraux et périphériques, *C. R. Soc. Biol.*, Paris, 1936, **122**, 57–59.
6. Drabovitch, W., La formation des réflexes conditionnés et la chronaxie, *Encéphale*, 1937, **32**, 93–99.
7. Fisher, R. A., *Statistical methods for research workers*, Edinburgh: Oliver & Boyd, 1936.
8. Hilgard, E. R., & Marquis, D. G., *Conditioning and learning*, New York: D. Appleton-Century Co., 1940.
9. Kellogg, W. N., & Walker, E. L., An analysis of the bilateral transfer of conditioning in dogs, in terms of the frequency, amplitude, and latency of the responses, *J. gen. Psychol.*, 1938, **18**, 253–265.
10. Partridge, G. E., Experiments upon the control of the reflex wink, *Amer. J. Psychol.*, 1899, **11**, 244–250.
11. Prosser, C. L., & Hunter, W. S., The extinction of startle responses and spinal reflexes in the white rat, *Amer. J. Physiol.*, 1936, **117**, 609–618.
12. Schäfer, E. A., Note on the occurrence of ganglion cells in the anterior roots of the cat's spinal nerves, *Proc. Roy. Soc.*, 1881, **31**, 348.
13. Sherrington, C. S., *Integrative action of the nervous system*, New Haven: Yale Univ. Press, 1906.
14. Shurrager, P. S., Further observations on 'conditioning' and 'extinction' in the spinal dog, *Psychol. Bull.*, 1939, **36**, 660.

15. ——, & CULLER, E., Conditioning in the spinal dog, *J. exp. Psychol.*, 1940, **26**, 133–159.
16. ——, & CULLER, E. A., Phenomena allied to conditioning in the spinal dog, *Amer. J. Physiol.*, 1938, **123**, 186–187.
17. ——, & ——, Reflex phenomena in the delimited spinal cord of the dog. *Amer. J. Physiol.*, 1939, **126**, 630.
18. ——, & SHURRAGER, H. C., Further analysis of central or synaptic learning and the peripheral response in the spinal dog, *Psychol. Bull.*, 1940, **37**. (Title in press.)
19. STUDENT, New tables for testing the significance of observations, *Metron*, 1925, **5**, 105–108.
20. WINDLE, W. F., Neurones of the sensory type in the ventral roots of man and of other mammals, *Arch. Neurol. Psychiat.*, Chicago, 1931, **26**, 791–800.

THE INFLUENCE OF SUCCESS AND FAILURE ON THE RESUMPTION OF AN INTERRUPTED TASK[1]

BY HELEN HOWARD NOWLIS

Yale University

I. Introduction

The problem of discovering the factors which determine the resumption of an interrupted task has developed considerable importance in recent years because of its systematic relation to the concept of substitute activity. This latter concept refers to the commonly observed fact that when one kind of activity or reward is prevented from occurring some other activity or reward can in some respects have the same effects on the organism's future behavior as the first would have had if it had occurred. For example, under normal circumstances a child may fret or cry if a toy with which he is playing is taken away from him. Return of the toy will usually stop the crying. If a different toy is given him, however, there is some probability that his crying will stop as quickly as if the first toy had been returned and he will cease his efforts to secure the first one.

From a systematic standpoint this implies that the second response reduces the strength of motivation to the first; *i.e.*, playing with the second or substitute toy has reduced the strength of the child's desire (as measured by efforts) to play with the first. Any attempt to discover the conditions under which the second response has this effect requires a method of measurement of the strength of motivation to carry out the first activity or task.

It was to secure such measurements that Ovsiankina (7) elaborated the interrupted-task technique devised by Zeigarnik (11). A subject is given a task to do, the motivation for which is intrinsic to the task itself or derives from the subject's desire for the experimenter's approval. At some point in the progress of the task, but before it is completed, the activity is interrupted. After an interval of relative freedom or of occupation with another task, the subject is given an opportunity to resume the first task.

The measure of strength of motivation to continue the first (interrupted) activity is based on the presence or absence of resumption of

[1] This study was made in the Institute of Human Relations; the present report is an abridgment of a dissertation presented to the Faculty of the Graduate School of Yale University in candidacy for the degree of Doctor of Philosophy in Psychology. The writer wishes to express her sincere appreciation of the many invaluable suggestions which have been made throughout this investigation by Professor Robert R. Sears.

that activity. It is assumed that at the time of interruption there was still a measurable amount of motivation to continue the task and that if this amount suffers no reduction there will be a resumption when opportunity is offered. These assumptions are supported by the work of Lissner (5), Mahler (6), and Ovsiankina (7). In general it has been shown that the interruption is followed by resumption whenever no intervening activity which might operate as a substitute occurs.

The measurement of any reduction in the strength of motivation to continue the first activity can be derived from a comparison of the frequency of resumption without intervening activity with frequency following any activity the substitutive qualities of which it is desired to test. It is clear that with the interrupted-task technique the strength of motivation to the original activity is operationally defined in terms of the frequency of resumption of that task, and that the measure of reduction of the strength of the motivation involves a comparison of the frequency of resumption following sheer interruption with frequency following the activity which is supposed to have reduced the strength. *A smaller frequency of resumption with the latter condition indicates that the intervening activity has served as a substitute activity.*

Analysis of the interrupted task situation shows that there are essentially four facts which seem to be of importance. The individual accepts a given task and, if he is not interrupted, will usually carry the task through to its conclusion. Once the conclusion is reached, he will cease working on that task. If, however, he is interrupted at some point in the series of acts leading to the completion of the task, he may object to being interrupted, even refuse, and if later given an opportunity will often resume and complete the activity. The interpolation of another activity following the interruption of the original one may result in a marked reduction in the resumption of the interrupted activity.

From the point of view of reaction psychology [2] the first fact to be accounted for is that the individual will continue to make a series of responses until he reaches a certain point and beyond this point will cease to make the responses. The continuation of action is dependent on the existence of *instigators*, which are defined as *the specified antecedent conditions of which any predicted response is the consequence.* There may be as many instigators to a response as there are antecedents which can be specified, and, unless there is evidence to the contrary, it is assumed that they represent the total amount (or strength) of instigation to that response.

[2] For a statement of this point of view see Dollard, Doob, Miller, Mowrer and Sears (1). The position adopted in the present discussion, together with certain concepts and their definitions, are those which have been applied to a variety of problems by these authors.

A second concept arises from the fact that, beyond a certain point, the individual will cease to make the responses which he has been making. The final response in a sequence beyond which the individual will cease to carry on the sequence of activities is a *goal response:* a goal response is *that response which reduces the strength of the instigation which produced it.*

In the present experiment, completion of a task in accord with the experimenter's instructions represents a goal response. In addition to *instigation-to-completion* there is other instigation involved in the interrupted-task situation. Almost everything that occurs even semi-officially in American college life is competitive. The mere mention of the word *test*, especially if it is related to psychology, makes the subject feel sure that, as on similar occasions in the past, his performance will be rated with reference to that of his fellow students. He wants his performance to compare well with that of the others. It is reasonable to suppose, therefore, that in addition to instigation to which the reaction to completion is a goal response there may be *instigation to which the reaction to success is a goal response.* Completion in itself may be considered a sort of success, but whether or not it functions as success for the subject depends on a number of factors of which an important one is its relation to his momentary level of aspiration.

What is perhaps the reverse of instigation to the goal response of reaction to success may also be operative in the interrupted task situation, at least as it occurs in the present investigation. A subject knows from past experience that a test situation offers at least two possibilities, that of showing how well he can do and that of showing how poorly he can do. He knows that two courses are open to him. He can attempt the solution of a problem with the chance of attaining success or he can avoid the problem and thus avoid the chance of doing poorly. Since resumption is his own choice and there is no external pressure in either direction he is free to choose resumption and possible success or non-resumption and certain avoidance of failure. A third source of instigation, therefore, may be postulated, *instigation to which the reaction to avoidance of failure constitutes a goal response.*

That there may be a marked reduction in the resumption of the interrupted activity when a second activity follows the interruption, leads to the definition of *substitute response* as *a response which reduces the instigation to a goal response which is specified as having been prevented from occurring.* A substitute response has *substitute value* to the extent to which it reduces the instigation to the original response.

A number of previous investigators have shown several factors to be important in determining the size of this value.

1. The nature of the original activity: resumption is more frequent with a task which has a definite goal than with one which lacks an objective goal and is more or less continuous (Ovsiankina, **7**; Zeigarnik, **11**).

2. The point at which interruption occurs in the progress of the original activity: interruption at points immediately after the instructions have been given and just before the completion of the task produces the highest frequency of resumption (Katz, **3**; Ovsiankina, **7**; Zeigarnik, **11**).

3. The nature of the interruption: 'accidental' interruptions produce more frequent resumption than do those which are clearly intentional on the part of the experimenter (Lissner, **5**).

4. The intensity of the primary action tendency: resumption is less frequent when the original task has been performed to satiation (Sliosberg, **10**).

5. Four characteristics of the period following the interruption: (*a*) the degree of similarity between the interrupted and the second activity (Lissner, **5**); (*b*) the degree of difficulty of the second activity relative to the first (Lissner, **5**); (*c*) the degree of contact between the tension systems involved in the two tasks (Lissner, **5**); and (*d*) the degree of reality of the second activity (Sliosberg, **10**; Mahler, **6**).

In addition to the above characteristics of the period following the interruption, there is some indication that, within wide limits, the mere amount of time elapsing between the interruption and the opportunity to resume is important.

Accepting, for the present, the findings of these earlier studies by the Lewin group as a point of departure, but divorcing them as far as possible from their relation to Lewin's theory of substitute activity (**4**), the present investigation proposes to hold all of these factors constant [3] and introduce success and failure on both the original and the second activity as additional factors which may influence resumption of the interrupted activity. Two main problems have been investigated: (1) the effect on the subsequent resumption of an interrupted activity of superimposing success or failure on the interruption of that activity, and (2) the effect on the resumption of an interrupted activity of success and failure superimposed on the completion of a second activity.

The reactions to success and to avoidance of failure are two of the three goal responses mentioned above as being involved in the interrupted-task experimental procedure. From a systematic standpoint, therefore, the present problem has been essentially that of determining the effect on resumption of the occurrence or non-occurrence

[3] With the possible exception of *dynamic contact*. For a discussion of this point, see Section IV.

of these three goal responses both at the point of interruption of the first activity and superimposed on the completion of the second.

From the work of the Lewin group there are several suggestions that feelings of success or failure on the part of the subject may influence resumption.

In discussing the results of her experiment on interrupted activity, Ovsiankina (7) suggests that for 'ambitious' subjects the experimental situation represents an occasion for testing their ability and if the task promises success there will be resumption, while there will be avoidance if it promises failure. Zeigarnik (11) has shown that the tension systems corresponding to an interrupted activity may be destroyed by sufficiently strong variations of tension in the whole person and in this connection mentions affective variations produced naturally or artificially, suggesting that strong feelings of failure may be among these. Lissner (5) has shown that difficulty of the second activity is likewise an important determiner of substitute value. The relation of the concept of 'difficulty' to those of 'success' and 'failure,' as it applies to the interpretation of the results of the present study, will be considered in detail in Section IV.

II. Method

The method of the present investigation is, with certain exceptions, essentially that of the interrupted task as developed by Ovsiankina (7) and elaborated by Lissner (5). In general, all those properties of the tasks which have been found by the Lewin group to affect the resumption of an interrupted activity were held constant by using the same tasks for each subject. Intensity of need was presumably held constant by using equivalent groups of twenty subjects in each experimental group. Tasks were chosen in which individual differences in ability would not produce too great variation in performance and which would, nevertheless, allow success and failure in the tasks to have some reality. The procedure for each subject was kept as rigid as seemed consistent with the best operation of the experimental variables. At no point, however, were adequate conditions for success and failure sacrificed intentionally for strict rigidity of experimental procedure. Variations in the response of individuals to praise and reproach determined the amount and, to some extent, the kind of both.

1. *The Tasks Used*

In selecting the tasks to be used it seemed wise to choose as the original or interrupted task one which met the following requirements: (*a*) it should have a well-defined goal and the responses made during the progress of the activity should have a definite relation to that goal;[4] (*b*) it should be capable of being interrupted at some specific point; (*c*) it should be interesting enough to the subject so that he will not welcome the interruption; and (*d*) it should allow success or failure without completion. The second activity should have a substitute value near 50 percent under neutral conditions [*i.e.*, neither success nor failure stimulation] so as to allow fluctuations in both directions under experimental conditions of success and failure. Neither activity should be highly preferred over the other.

Because it is almost impossible to predict *a priori* what activities or tasks will best fulfil these requirements, a series of six different tasks, which included cancelation, simple addition, card sorting, dart throwing, a ball-in-wheel puzzle and a jig-saw puzzle, were presented in various com-

[4] Both Zeigarnik (11) and Ovsiankina (7) have shown that the greatest resumption is obtained with tasks which have a well-defined goal, in their words 'Endhandlungen' as contrasted with 'fortlaufende Handlungen.'

binations, to a small group of subjects. The combination of the ball-in-wheel puzzle followed by the jig-saw puzzle was the most satisfactory from the standpoint of the criteria mentioned above.

The ball-in-wheel puzzle chosen as the original or interrupted activity consisted of a glass-covered box in which there were a wheel with eight cogs in it and eight small balls, four black and four white. The cogwheel, which turned freely on its shaft, was surrounded by a fixed cardboard ring in which there was one small opening through which the balls might pass. The subject was required to put the balls in the cogs of the wheel, alternately a black and a white. The average time required by the 180 subjects in the main experimental groups to put in six of the eight balls was five minutes.

The second task was a jig-saw puzzle cut from one piece of 3-ply wood, 12 inches by 9.5 inches, and stained with a dark stain. There were 15 pieces, ten of which had at least one right angle. The grain showed through the stain and was the best cue for putting the puzzle together. The average time required for completion by the nine experimental groups was ten minutes.

2. *Subjects*

The subjects for the main part of the experiment were one hundred and eighty undergraduate men. One half of this number were registered in elementary psychology courses and served in this experiment as a part of a course requirement. The other ninety were obtained through the University Student Employment Office and were paid fifty cents for their services.

Each of the nine experimental groups was composed equally of paid and unpaid subjects. In no respect were there any apparent differences in the performance of the two groups of subjects. They were equally coöperative and seemed equally well motivated once they were in the experimental situation. In several instances the paid subjects would even start to leave before they had been paid.

In addition to the above one hundred and eighty, there were thirty subjects who served in control groups. Since there seemed to be no differences between the two groups, and the paid subjects were more readily available, the control groups consisted of paid subjects only.

So far as could be determined none of the subjects knew or suspected the purpose of the experiment. They were told at the time of making their appointments that they were to serve in an experiment on perception and would do a number of puzzles. At the end of the session the subject was asked what ability or function he felt the experiment might adequately test. The answers almost invariably were perception, learning or mechanical ability. In only three instances was there any approach to the true purpose of the experiment and the data for these subjects were discarded. At no time were the subjects told the real purpose.

3. *Procedure*

The subjects were divided into nine experimental groups of twenty subjects each. These groups represented all nine possible combinations of success, neutral and failure conditions on both the first and the second tasks. For convenience they are represented in Table I.

TABLE I

THE NINE EXPERIMENTAL GROUPS

Task I	Task II		
	Success	Neutral	Failure
Success	*A* Success I–Success II	*B* Success I–Neutral II	*C* Success I–Failure II
Neutral	*D* Neutral I–Success II	*E* Neutral I–Neutral II	*F* Neutral I–Failure II
Failure	*G* Failure I–Success II	*H* Failure I–Neutral II	*I* Failure I–Failure II

With each of the nine groups Task I was always interrupted and Task II was always completed. Success and failure were produced by statements of the experimenter at the point of interruption of Task I and at the conclusion of Task II.

Groups A, B and C (Success I) [5] were interrupted on the first task but were made to feel that they had done very well; Groups D, E and F (Neutral I) were merely interrupted on the first task by being requested to help the experimenter check a list of numbers for a minute; Groups G, H and I (Failure I) were interrupted and made to feel that they had done very poorly in terms of their speed. For Task II, Groups A, D and G (Success II) completed the second task and were highly praised; to Groups B, E and H (Neutral II) no comment was made when the subject completed the second task; Groups C, F and I (Failure II) completed the second task but were told that they had been unusually slow.[6]

In all cases the subject was seated at a desk on which had been placed the cogwheel and the jig-saw puzzles. The cogwheel was placed in front of him with the following instructions:

> "The first thing is this. There are four black balls and four white ones. You are to put the balls in the cogs of the wheel, alternately a black and a white. Do you understand? Ready, go."

The experimenter sat in a chair near the desk and busied herself with a sheaf of papers. It was not difficult to follow very unobtrusively the progress of the subject, for after each ball was in a cog of the wheel it was necessary to tap the puzzle to make the wheel turn so that another cog was opposite the hole in the cardboard ring. In most instances the 'ah's' and 'there's' of the subject also signified the correct placement of a ball. When six balls were correctly placed the subject was interrupted with a Success, Neutral or Failure interruption.

Following the interruption, the jig-saw puzzle was pushed forward on the desk by the experimenter with the following remarks:

> "The jig-saw puzzle makes a rectangle. It is all cut out of one piece of wood so that the grain matches, and the long piece is one dimension."

At this point the experimenter retired to a desk at the opposite end of the room, facing the one at which the subject was working. She again busied herself with books, papers and writing and apparently ignored the subject completely. When the jig-saw was completed, it was necessary for the subject to call the attention of the experimenter to the fact. At this time the experimenter either praised, reproved, or said casually, "Just a second, please," and continued writing hurriedly for a period of two minutes.[7] During this period the subject was free to resume the interrupted activity if he chose. A detailed record was made of his behavior, including the point at which he resumed. If he did resume he was allowed to work until he had completed the task or put it aside voluntarily. If he did not resume spontaneously he was asked, at the end of the two-minute period, to choose one of the two tasks to do again.

At the conclusion of the experiment the subject was asked the following questions:

1. What ability or function do you think this experiment might adequately test?
2. Which type of puzzle do you prefer doing?
3. Which puzzle do you think is easier?
4. Did you have a theory or a working hypothesis when you were doing the cogwheel puzzle?

4. *Success, Neutral and Failure Conditions*

The problem of making the subject feel that he really had succeeded or failed on either of the tasks was a difficult one. Because it seemed advisable to keep the average time spent on the tasks as constant as possible for each group, all those subjects who actually did well could not be put in the Success groups and those who did poorly in the Failure groups.[8] With a few exceptions,

---

[5] The Roman numerals I and II following Success, Failure or Neutral refer to Task I and Task II. Thus Success I means Success stimulation accompanying the interruption of Task I. Success I-Success II refers to Group A, Success I-Neutral II refers to Group B, etc.

[6] The Success, Neutral and Failure conditions are discussed more fully in Section 4 below.

[7] It has been shown by Ovsiankina (7) that spontaneous resumption, if it occurs at all, will usually occur within 4 to 20 seconds. The 2-minute interval was decided on as having sufficient margin to allow for any resumptions. There is some indication from the present data that the 20-second interval specified by Ovsiankina is not sufficient.

[8] Ovsiankina reports some evidence to the effect that, within 4 to 20 minutes, the mere time elapsing after interruption does not greatly influence resumption.

the subjects were allocated to the nine groups in whatever order they appeared at the laboratory. However, if a subject scheduled for Success on the first task was obviously doing very poorly and seemed to feel that he was not doing well he was shifted to a Neutral or Failure group.

In general, the procedure for producing feelings of success and failure was that used by Sears (8). Comments and attitudes of the experimenter were used to supplement the reporting of false time scores. For the Neutral condition no reference was made to the time score and evaluative comments were avoided.

For Success and Failure on the first task the experimenter interrupted the subject just after the sixth ball had been put in place and either praised or reproached. In the case of Neutral on the first task the experimenter said, without looking up, "Say, would you do something for me?" and, handing the subject a sheet of numbers, "Just read these totals to me while I check them." The average time for reading these numbers was 50 seconds, approximately the time consumed by the Success and Failure procedures. For Success and Failure on the second task the appropriate comments were made when the task was completed. When the second task was Neutral the experimenter did not even look up from her work but said, "Just a minute," and continued writing rapidly for the two-minute period.

A subjective estimate of the response of the subject to interruption and to Success and Failure stimulation was made on a four-point scale. Approximately 80 percent of the subjects in each group made some observable response to these conditions.

5. *Control Groups*

Since it became increasingly evident as the experiment progressed that the time spent on the second task would range from four to twenty minutes and that the average time would be from three to five minutes longer for the Failure than for the Success groups, three additional groups of ten subjects each were used to determine the effect of varying amounts of time elapsing between interruption and the opportunity to resume. In all of these groups the interruption of the first task was Neutral and the second activity was one which had no possible relation to the interrupted activity. The only variable was the amount of time consumed by the second activity.

The subjects of the first group (Neutral-1 minute) were given the cogwheel puzzle with the usual instructions and were interrupted after placement of the sixth ball in the same way in which the Neutral I groups in the main experiment were interrupted. They were asked, as a favor to the experimenter, to read the same list of totals while she checked them. When the figures had been read the experimenter said, "Thanks," and continued working with the figures for two minutes while the subject was free to resume the cogwheel.

The second group (Neutral-5 minute) were given the cogwheel as usual and when six balls were in the experimenter said, "Say, would you do something for me? I promised one of the professors that I would put an assignment on the board in Room 104 for his seminar and I forgot it completely. Would you be willing to go downstairs and write this assignment on the board? I think you can read it all right." The experimenter then gave the subject a piece of paper on which was written a long German reference and read it over with him to be sure he understood it. The subject had to go from the third floor to the first floor, write the reference on the blackboard and return to the experimental room. This routing occupied an average of five minutes for the ten subjects. When he returned to the experimental room the experimenter was busily occupied and greeted his arrival with only a casual "Thanks." She continued working for the usual two-minute period and noted carefully the behavior of the subject. If, at the end of the two minutes, he had not resumed he was asked whether he would prefer to work on the cogwheel or do a jig-saw puzzle.

The procedure for the third group (Neutral—15 minute) was essentially the same except that the subject was given a paper with four long references on it. The average time for this group was fifteen minutes.

The average interpolated time for the twenty subjects of the combined Neutral-5 minute and Neutral-15 minute groups closely approximated that required for the interpolated (second) activity of the Neutral I-Neutral II group of the main experiment (10 minutes). The frequency of resumption in this control group provided a basal resumption score which, when compared with the frequency of resumption in the Neutral I-Neutral II group, would permit an estimate of the basic substitute value of the jig-saw puzzle.

III. Results

In Table II is presented the total number of subjects in each group who resumed Task I. Three types of resumption could be distinguished and separate figures are given for each type. These are (1) *resumption with completion*, in which the subject spontaneously resumed the interrupted task within the two-minute free period and continued working until he had completed it; (2) *resumption without completion*,[9] in which the subject spontaneously resumed the interrupted task within the two-minute limit but put the cogwheel aside without having completed it; and (3) *choice resumption*, in which the subject did not resume during the two-minute period but, when asked to choose one of the tasks at the end of that period, chose the interrupted task.[10]

TABLE II

Resumption of the Interrupted Task Under the Nine Experimental Conditions

Group *	Resumption with Completion		Resumption without Completion		Choice Resumption		Total	
	N *	%	N	%	N	%	N	%
Success I–Success II	7	35	1	05	4	20	12	60
Success I–Neutral II	8	40	1	05	5	25	14	70
Success I–Failure II	3	15	8	40	3	15	14	70
Neutral I–Success II	15	75	1	05	3	15	19	95
Neutral I–Neutral II	5	25	3	15	7	35	15	75
Neutral I–Failure II	8	40	0	00	3	15	11	55
Failure I–Success II	9	45	4	20	4	20	17	85
Failure I–Neutral II	4	20	4	20	3	15	11	55
Failure I–Failure II	3	15	6	30	5	25	14	70
Combined Success I	18	30	10	17	12	20	40	67
Combined Neutral I	20	47	4	07	13	22	45	76
Combined Failure I	16	27	14	23	12	20	42	70
Combined Success II	31	52	6	10	11	18	48	80
Combined Neutral II	17	28	8	13	15	25	40	66
Combined Failure II	14	23	14	23	11	18	39	64

* N for each single group is 20, for each combined group, 60.

Examination of Table II indicates that, for the total resumption scores, the only difference between Success, Neutral and Failure conditions in either task which has a *P*-value [11] of less than .04 is that

[9] *Cf.* Ovsiankina (7) *tendency to resume* and Lissner (5) *playful resumption* and *resumption without completion.*

[10] This is almost equivalent to Mahler's (6) *resumption after instruction* except that, in this case, all those subjects who had not resumed to completion were required to make a choice between doing Task I or Task II.

[11] *P* is the chances in 100 that the differences between two sets of data can be attributed to chance. Throughout this discussion *P* is derived from the Critical Ratio of the difference between two percents.

between Success and Failure on Task II, Task I being Neutral. Similar comparisons for the combined groups [12] (Table II) show no differences which would occur by chance in less than four cases in 100.

However, if one considers only resumption with completion, as there may be some justification for doing,[13] differences in the effect of Success and Failure in both the first and the second tasks become more apparent. The following analyses will be based on the figures for *resumption with completion.*

1. *The Effect of Success and Failure Stimulation in the First Task on Its Resumption*

The results for Success and Failure stimulation in the first task are, in general, inconclusive (Table III). A comparison of the Success I–Neutral II group with the Failure I–Neutral II group indicates that there is a tendency in the direction of increased resumption with Success stimulation on the first task. The difference between the two groups is 20 percent (P = .08). The difference between the combined Success I group and the combined Failure I group is only 3 percent in the direction of decreased resumption under Failure conditions (P = .36). The combined Success I and the combined Failure I groups both differ in the direction of decreased resumption from the combined Neutral I group with differences of 17 percent and 20 percent, respectively (P = .03 and .01).

TABLE III

COMPARISON OF RESUMPTION UNDER SUCCESS, NEUTRAL AND FAILURE STIMULATION IN TASK I

Group	Difference (%)	Critical Ratio	P
Success I–Neutral II—Failure I–Neutral II	20	1.41	.08
Success I–Neutral II—Neutral I–Neutral II	15	1.02	.15
Failure I–Neutral II—Neutral I–Neutral II	15	.38	.36
Combined Success I—Combined Failure I	03	.36	.36
Combined Success I—Combined Neutral I	17	1.95	.03
Combined Failure I—Combined Neutral I	20	2.32	.01

[12] The combined Success I group is obtained by adding the results for the three groups in which Success I appeared as a variable. The same procedure was followed in the other combined groups.

[13] A rank-order correlation between *total resumption* and *resumption with completion* is +.44. Since one of the goal responses postulated in Section I is that of reaction to completion, we are interested in the strength of the instigation to complete Task I after the completion of Task II. *Resumption with completion* would seem to be the critical measure for this. Resumption without completion in most instances was very brief, not exceeding one minute in any case, and indicates little about the strength of instigation to complete Task I after the completion of Task II.

2. *The Effect of Success and Failure Stimulation in the Second Task on the Resumption of the Interrupted Task*

The effects of Success and Failure stimulation following the completion of the second task are much more clearly differentiated (Table IV). A comparison of the Neutral I–Success II group with the

TABLE IV

COMPARISON OF RESUMPTION UNDER SUCCESS, NEUTRAL AND FAILURE STIMULATION IN TASK II

Group	Difference (%)	Critical Ratio	*P*
Neutral I–Sucess II—Neutral I–Failure II	35	2.39	.01
Neutral I–Success II—Neutral I–Neutral II	50	3.64	<.01
Neutral I–Failure II—Neutral I–Neutral II	15	1.02	.16
Combined Success II—Combined Failure II	29	3.04	<.01
Combined Success II—Combined Neutral II	24	2.75	<.01
Combined Failure II—Combined Neutral II	05	.63	.27

Neutral I–Failure II group shows that resumption under Success II conditions is greater by 35 percent than resumption under Failure II conditions ($P = .01$). A similar comparison of the Combined Success II groups with the Combined Failure II indicates a difference of 29 percent ($P = <.01$) in the same direction.

Although the Combined Success II group differs from the Combined Neutral group by 24 percent in the direction of increased resumption ($P = <.01$), the Combined Failure II group is strikingly like the Combined Neutral II group. The difference is only 5 percent ($P = .27$) with resumption slightly higher in the Combined Neutral II group.

3. *Analysis of Variance*

Since Success and Failure stimulation following the second task have been shown to be highly influential in determining the amount of resumption of the interrupted first task, it is possible that the second task is to some extent masking the effects of Success and Failure stimulation on the first task. It is possible, by using the Fisher analysis of variance, to determine the relative importance of varying degrees of success [14] in the two tasks in determining the total variance. The method is essentially a test of significance which allows for the simultaneous comparison of the variables involved. The total variation is divided into parts corresponding to the degrees of freedom allowed by each variable; the amounts of variation contributed by each variable to the total are compared.

[14] Success, Neutral and Failure may be considered three degrees of success with Success the maximum and Failure the minimum.

The analysis of the variance in the present data, following the method of Snedecor (9), is summarized in Table V. In order to

TABLE V

SUMMARY OF THE ANALYSIS OF VARIANCE FOR RESUMPTION WITH COMPLETION

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F *
Total	17	72.5		
Degree of Success on Task I	2	13.8	6.9	4.06
Degree of Success on Task II	2	27.5	13.8	8.12
Interactions	4	16.2	4.1	2.41
Replication†	9	15.0	1.7	

* F is the ratio between the mean square of one source of variation and the variation used as an estimate of error. It is an adaptation by Snedecor (9) of Fisher's z (2), which in turn is an extension of t appropriate to cases in which more than two variables are being compared. It is derived from the theory of errors and is less affected than t by obtained deviations from normality. The P-values of F are given in a table of F (Snedecor, p. 184).

† For the nine groups a rank-order correlation of the odd-even subjects is + .62.

increase the possible number of degrees of freedom and to have a known variable, other than the degree of success on the two tasks, as a basis for the comparison, each of the experimental groups was divided into two parts by taking the odd and even subjects from each of the original groups. This provides two measures for each of the nine main conditions instead of one and permits a statement of the influence of degree of success in terms of how much more the various degrees contribute to the total variance than the sampling error contributes.

In the analysis of variance (Table V), the F-values for degree of Success on Task I and on Task II are 4.06 ($P = < .05$) and 8.12 ($P = < .01$) respectively. The amount of variance contributed to the total by the degree of success (the combined Success, Neutral and Failure conditions) on either task is from five to eight times that contributed by replication. In view of the fairly low reliability of the odd and even groups the significance of variations in degree of success on either task is considerable.

One further fact gained from this analysis is relevant to the question of the possible blanketing of the effect of variations in degree of Success in Task I. Task I does contribute to the total variance but the contribution of Task II is approximately twice that of Task I. It is also indicated that there is an interaction between the conditions of Task I and those of Task II which might further account for the small differences obtained between the Success I and Failure I groups when direct comparisons of these groups are made.

The F-value for interaction falls so far short of a P-value of .05 that it is hardly worth pursuing further in an analysis of covariance.

On the basis of these results it would seem justifiable to conclude that the degree of success in the interrupted activity is somewhat influential in determining the resumption or non-resumption of that activity, but that the significantly greater influence of degree of success in the second task may overshadow it.

4. *The Influence of Success and Failure Stimulation on Subsequently Expressed Preferences for the Two Tasks*

At the conclusion of the experiment each subject was asked which of the two tasks he preferred. These preferences are summarized in Table VI.

TABLE VI

PREFERENCES FOR AND JUDGMENTS OF DIFFICULTY OF THE TWO TASKS EXPRESSED AT THE CONCLUSION OF THE EXPERIMENTAL SESSION

Group	Prefer Task II		Task II easier	
	No.	%	No.	%
Success I–Success II	14	70	9	45
Success I–Neutral II	10	50	11	55
Success I–Failure II	7	35	4	20
Neutral I–Success II	13	65	14	70
Neutral I–Neutral II	10	50	15	75
Neutral I–Failure II	8	40	8	40
Failure I–Success II	14	70	15	75
Failure I–Neutral II	15	75	15	75
Failure I–Failure II	12	60	16	80

Regardless of whether the Success or Failure was on the interrupted or the completed task, the one which was accompanied by Success stimulation was preferred. A comparison of the percent of subjects in the Success I–Failure II group who preferred Task I with the percent of subjects in the Failure I–Success II group preferring Task I gives a difference of 35 percent ($P = < .01$) in the direction of preference for the task in which there was Success stimulation. It is interesting to note here that in the Success I–Neutral II group 50 percent of the subjects preferred Task II, while in the Success I–Success II group 70 percent of the subjects preferred Task II. This might indicate that Neutral II following Success is more like Failure in Task II than is Success II following Success I. A similar state of affairs seems to hold when the Failure I–Neutral II and the Failure I–Failure II groups are compared. The preferences for Task II are 75 percent and 60 percent, respectively. There seems, therefore, to be some indication that Neutral following Success is experienced by the subject as a mild failure, while Neutral following Failure is ex-

perienced by the subject as definite Success. A comparison of Success following Failure (70 percent preferring II) and Neutral following Failure (75 percent preferring II) further substantiated the latter conclusion.

In no group was a Failure task preferred over one in which there was either Success or Neutral stimulation.

An analysis of variance based on the number in each group preferring Task I [15] indicates that degree of Success on the second task is more influential than degree of Success on the first but the *P*-values corresponding to *F*-values of 21.5 and 3.15 for nine and two degrees of freedom are considerably greater than .05.

5. *The Influence of Success and Failure Stimulation on Subsequently Expressed Judgments of the Relative Difficulty of the Two Tasks*

The subjects were asked also, at the conclusion of the experimental session, which of the two tasks they considered the easier. The results of these judgments are also summarized in Table VI.

A similar trend to that found in the case of preference is shown, although in this instance the results are not quite so consistent. A comparison of the percent of the subjects in the Success I–Failure II group who judged Task I as the easier with the percent of the Failure I–Success II group judging Task I the easier gives a difference of 55 percent in the direction of more judging easier the task on which there was Success stimulation ($P = < .01$).

An analysis of variance for these data indicates that the degree of Success on the first task is more than twice as influential as degree of Success on the second and that both contribute more to the total variance than does a sampling error.[16] The *P*-values corresponding to *F*-values of 9.2 and 3.5 for nine and two degrees of freedom are less than .01 and considerably more than .05, respectively.

6. *The Effect of the Amount of Time Elapsing between the Interruption and the Opportunity for Resumption*

Although the average time spent on Task II was 10.56 minutes for the nine experimental groups, the range was from three to twenty-five minutes and the averages for the individual groups were from 9.3 to 14.1 minutes. Because the Failure groups were generally slower than the Success groups, some check on the influence of the length of time elapsing between interruption and resumption was necessary.

In Table VII are presented the number of resumptions in the three control groups, Neutral–1 minute, Neutral–5 minutes and

[15] For the nine groups a rank-order correlation of the odd-even subjects is +.43.

[16] For the nine groups a rank-order correlation of the odd-even subjects is +.69.

TABLE VII

RESUMPTION OF THE INTERRUPTED TASK IN GROUPS WITHOUT SUBSTITUTE ACTIVITY

Group	N	No. resuming	% resuming
Neutral– 1 minute	10	8	80
Neutral– 5 minutes	10	8	80
Neutral–15 minutes	10	6	60

Neutral–15 minutes. In these three groups the factor of elapsed time is not complicated by an activity which might be a substitute for the first task. There is no difference in resumption between the 1- and the 5-minute groups, and the difference of 20 percent ($P = .16$) between these and the 15-minute group, although indicating a tendency in the direction of reduced resumption, is not very significant in comparison with the differences obtained as the result of the main experimental variables. It must be noted, however, that the number of subjects in these groups is small and the conclusion must be correspondingly tentative.

Additional evidence in support of this conclusion is offered by a further analysis of the data of the nine main experimental groups. Each group may be divided into two parts on the basis of the time spent on Task II and the ten subjects who spent the longest time on Task II may be compared with the ten subjects who spent the shortest time. The results of this analysis are presented in Table VIII.

TABLE VIII

COMPARISON OF RESUMPTION FOR THE HALF OF EACH GROUP WHICH SPENT THE LONGER TIME ON TASK II AND THE HALF WHICH SPENT THE SHORTER TIME

Group	Longer half		Shorter half	
	Av. time	No. resuming	Av. time	No. resuming
Success I–Success II	10.2	4	5.3	3
Success I–Neutral II	12.0	1	4.6	7
Success I–Failure II	12.0	2	6.8	1
Neutral I–Success II	13.0	7	5.7	8
Neutral I–Neutral II	17.3	3	9.5	2
Neutral I–Failure II	15.8	4	10.6	4
Failure I–Success II	11.7	6	5.8	3
Failure I–Neutral II	17.5	3	6.6	1
Failure I–Failure II	14.7	2	8.5	1
Average	13.9		7.0	
Total		32		30

The totals for the two groups indicate that despite a seven-minute difference in the average time of the longer half and the shorter half of each group, there is a difference in total resumption of only 2 percent.

On the basis of this combined evidence it would seem safe to conclude that the discrepancy between the times spent on Task II in the main experimental groups was not an important variable.

7. *The Substitute Value of Task II*

The basic substitute value of Task II is obtained by comparing the combined Neutral–5 minute and Neutral–15 minute groups, in which the interpolated activity had no relevance to the original task, with the Neutral I–Neutral II group where Task II occurred without either Success or Failure stimulation. This comparison shows a difference in resumption of 45 percent.

Lissner (**5**) has adopted the procedure of expressing substitute value in terms of the ratio between frequency of resumption without substitute activity and frequency of resumption with substitute activity. Expressed in these terms the substitute value of Task II under Neutral conditions is 2.8.

When Success and Failure are introduced as variables the differences in percent of resumption range from 5 percent more resumption in the Neutral I–Success II group (substitute value = .93) to 55 percent in the direction of less resumption in the Failure I–Failure II and the Success I–Failure II groups (substitute value = 4.6).

IV. Discussion

Two alternatives present themselves in the interpretation of these results with reference to Lewin's field theory of psychological forces. One is that the conditions produced by Success, Neutral and Failure stimulation can be equated to certain of the factors which have already been shown to influence resumption, *e.g.*, dynamic contact, difficulty or similarity. In such case these kinds of stimulation would be special instances of conceptualized variables which are already an integral part of the field theory. The other alternative is that the effects of Success, Neutral and Failure stimulation be accepted simply as a further set of empirically determined factors which influence the substitute value of an interpolated task.

In discussing *dynamic contact* in relation to substitute value Lewin states:

> We find very little substitute value if the two activities are psychologically *separated* through special circumstances of the situation. Such isolation can sometimes be realized by having the Experimenter say at the time of interruption: "Now we shall do an entirely new task . . . " (**7**, p. 185).

The experimental evidence on which he bases this statement is from Lissner (**5**, p. 40). Close examination of Lissner's data reveals that, for the group with *dynamic contact* (Im Zussamenhang), of the 43 percent who resumed, none resumed to completion, 29 resumed in

the sense that they did something other than the original task with the materials of the original task, and 14 percent resumed *in thought*. In the group without *dynamic contact* (Ohne Zussamenhang) 70 percent of a total of 85 percent resumptions were resumptions with completion. In other words, a high degree of dynamic contact reduces resumptions to a greater extent than a low degree and therefore may be said to provide the interpolated task with a greater substitute value.

In the present investigation Success and Failure stimulation at the point of interruption might be interpreted as breaking the dynamic contact between the two tasks. But the results, in terms of either resumption with completion or total resumption, do not support this suggestion. Both Success and Failure decreased the number of resumptions with completion to approximately 40 percent of the number obtained with a Neutral interruption. The reverse of the Lissner results is true. Therefore, if the effect of dynamic contact between two tasks is to decrease the number of resumptions of the interrupted task following the second task, then in the present instance either Success and Failure stimulation at the point of interruption *increased* the dynamic contact between the two tasks or some factor other than dynamic contact is operative.

The present results appear only at first sight to be in agreement with Lissner's finding that a relatively *difficult* substitute task decreases the resumption of an original interrupted task more than does a relatively easy one. Comparison of Success I–Failure II with Failure I–Success II and of Neutral I–Success II and Neutral I–Failure II suggests that the effects on resumption of Success, Neutral and Failure stimulation are dependent on the degree of perceived difficulty produced by these conditions. Careful examination of the data, however, shows that but four of the nine groups are used in these comparisons and they represent average relationships without regard to variability.

In the first column of Table VI has been given the percent of subjects in each group who judged the second task to be easier. The larger the percent for any group, therefore, the greater was the number of subjects who found the second task easier than the first. Since according to the Lissner results resumption should occur more frequently with easy substitute activity than with difficult, the groups having a large percent of 'easier' judgments should have a large percent of resumptions. But this proves not to be the case. When the nine groups are ranked in terms of the number of subjects in each group who found the second task easier, the rank-order correlation with the same groups ranked in terms of obtained resumption scores is −.08.

It appears, then, that identifying Failure stimulation with difficulty and Success stimulation with relative non-difficulty gives no predictive power whatever.

The factor of similarity appears, likewise, to permit of little differentiation between the groups on a theoretical basis, since the second task was equally similar to the first under all nine conditions. Despite dissimilarity of those factors which Lissner considers important determiners of similarity between two tasks, in the present problem the difference between resumption after an irrelevant second activity and resumption after the jig-saw under neutral conditions is of the same order as that obtained by Lissner between resumption without substitute activity and resumption following a task which was similar in terms of the material used, the type of activity involved and the finished product.

There is similarity, however, in the nature of the instigation to the two tasks. It was pointed out earlier that both tasks involved goal responses of *reaction to completion*, *reaction to success*, and *reaction to avoidance of failure*. So far as instigation is concerned, therefore, similarity is present to some degree and may be the factor which provides the second task with a basic substitute value under Neutral I–Neutral II conditions.

Since the resumption scores for the nine experimental conditions are not predictable on the basis of an identification of Success, Neutral and Failure stimulation with the conceptualized variables of *dynamic contact*, *difficulty* or *similarity*, the only alternative from the standpoint of field theory seems to be to adopt the suggestion of Zeigarnik (**14**) that sufficiently strong variations in affective state of the subject may destroy the tension system corresponding to the interrupted task and thus reduce resumption. Such an explanation would account for the low resumption scores in the Failure II groups but would not account for the fact that resumption in the Success II groups is significantly greater than in the Neutral II groups. It is necessary, therefore, to look elsewhere than to field theory for a systematic explanation of the results.

Inasmuch as a substitute response is one which reduces the instigation to an original interrupted activity and the strength of instigation is operationally defined in terms of frequency of resumption, the jig-saw puzzle does act as a substitute for the cogwheel. Under Neutral I–Neutral II conditions resumption of the interrupted activity was 25 percent as compared with 70 percent when the second activity was completely irrelevant to the interrupted task (Control group of combined Neutral–5 minute and Neutral–15 minute groups).

When the factors of success and failure are superimposed on the usual interrupted task situation, however, there is a wide variation

among the nine groups in terms of amount of resumption. The obtained values vary from 75 percent resumption under a Neutral I–Success II condition to 15 percent resumption for Success I–Failure II and Failure I–Failure II. In general, success on the second task favors resumption and failure militates against it (Combined Success II, 52 percent, Combined Failure II, 23 percent). This fact, taken in conjunction with the theoretical analysis of the experimental situation,[17] suggests that the effect of the occurrence on a second task of goal responses common to a first is to reinforce or strengthen the instigation to the first.

From this point of view the present results may be interpreted in the following way. Instigation has been defined as the specified antecedent conditions of which a predicted response is the consequence, and three components, *a*, instigation to which the reaction to completion is a goal response, *b*, instigation to which the reaction to success is a goal response, and *c*, instigation to which the reaction to the avoidance of failure is a goal response, have been specified. It can be assumed that the total strength of instigation is the combined effect of all the instigation represented by the specified goal responses in a given situation; therefore, the goal responses *a*, *b*, and *c* in combination represent the instigation to both Task I and Task II.

If the occurrence of a goal response on Task II generalizes to reinforce the instigation to Task I and goal responses *a*, *b*, and *c* in combination represent the instigation to Task I, the reinforcing value of the occurrence of these goal responses on Task II should be directly proportional to the number of goal responses which occur. It would, therefore, be predicted that under Success II conditions, where all three goal responses occur, reinforcement would be greater than under either Neutral II or Failure II conditions where only two (*a* and *c*) and one (*a*) goal responses, respectively, occur. This hypothesis is supported by the resumption scores for the combined groups (Success II, 52 percent; Neutral II, 28 percent; Failure II, 23 percent).

To account for the results of the nine experimental groups individually is more difficult. Since, by definition, a goal response is that reaction which immediately reduces the strength of instigation to a degree at which it no longer tends to produce the predicted response, the occurrence of any of the three goal responses will *reduce* the instigation to Task I and the amount of reduction will be directly proportional to the number of goal responses occurring. But the strength of instigation to Task I is not measured until after the completion of Task II and it is assumed that the instigation to a task reduced by the occurrence of a goal response to that task recovers

[17] *Cf.* Introduction.

partially when a second activity intervenes between the task and the point at which the instigation is measured. Therefore, the instigation to Task I reduced by the occurrence of any of the specified goal responses will partially *recover* by the time the strength of instigation to Task I is measured. Also, since the occurrence of a goal response reinforces the instigation to perform a task, the occurrence of any of the three goal responses on Task I will *reinforce* the instigation to Task I. As above, the occurrence of any of the three goal responses on Task II will generalize to reinforce the instigation to Task I. The strength of instigation to Task I at the time when it is measured will thus be a function of the occurrence on Task I or Task II, or by the interaction between Tasks I and II, of one, two, three, or none of goal responses *a*, *b*, and *c*.

Ideally, it should be possible to test this hypothesis, at least tentatively, by deducing the expected rank of each of the nine experimental groups and correlating the expected ranks with those actually obtained. If the relative importance of the three instigational components and a constant for the amount of recovery of instigation reduced by the occurrence of the goal responses were determined this could be done.[18]

It is difficult to consider the relation of this discussion to data obtained by other investigators for, in most instances reported in the literature, a variety of tasks was used in a single experimental session, the experiments were done in another culture, and knowledge concerning the manner in which subjects were obtained and the instructions given to them with regard to the general conditions of the experiment is not available. It is thus impossible to assume with any certainty the operation of the instigation components on which the present discussion has been based. It should be possible, in cases where the background of the subjects and the social factors arising from the experimental situation can be specified, to name the instigation components and consider the present hypothesis.

V. Summary

The interrupted-task technique was used for the study of the influence of success and failure on substitute activity; those characteristics of a task which have been shown by previous investigators to influence the resumption of an interrupted task were held constant. Verbally produced success and failure were introduced both

[18] Such a correlation, using a simple equal-weighting of the three components and one-half the value of a goal response for the recovery constant, was tried in the original dissertation (on file in Sterling Library, Yale University). A rank difference coefficient of +.76 was obtained. When this was corrected after the model of the formula for the coefficient of attenuation the value was +.98. Definitions, postulates, and detailed deductions are presented in the dissertation.

at the point of interruption of the first task and following the completion of a second task, and their effect on the subsequent completion of the interrupted task was determined. The findings, as obtained from 210 male undergraduates, may be summarized as follows:

1. Success and Failure stimulation on an interrupted task (Task I) do not produce significantly different amounts of resumption of that task following a second task. Both Success and Failure stimulation definitely reduce resumption in comparison with a Neutral interruption, the effect of Failure being slightly more pronounced than that of Success.

2. Success stimulation following the completion of a second task (Task II) increases resumption of the original activity to a significantly greater degree than does Failure stimulation or a Neutral condition. Failure stimulation and a Neutral condition on a second task are not significantly different in their effects on resumption of an original interrupted task.

3. Success, Neutral and Failure conditions on either the interrupted or the second task contribute significantly more to the total variance than does the sampling error. The comparatively greater effect of Success, Neutral and Failure conditions on the second task indicates that the results of similar variations on the interrupted task may be somewhat overshadowed.

4. Regardless of whether the interrupted or the completed task is considered, Success stimulation produces significantly more judgments of preference for a task than does Failure stimulation.

5. Regardless of whether the interrupted or the completed task is considered, Success stimulation produces significantly more judgments of 'easier' for a task than does Failure stimulation.

6. Within the limits of one and fifteen minutes, the time elapsing between the interruption of a first task and the subsequent opportunity for its resumption is not an important factor in determining resumption.

7. Variations in the substitute value of the same task, ranging from 4.6 to .93, were obtained when Success, Neutral or Failure conditions on the two tasks was the only variable.

These results are interpreted from the point of view of reaction psychology.

(Manuscript received October 8, 1940)

References

1. Dollard, J., Doob, L. W., Miller, N. E., Mowrer, O. H., Sears, R. R., Ford, C. S., Hovland, C. I., and Sollenberger, R. T., *Frustration and aggression*, New Haven: Yale Univ. Press, 1939. Pp. ix + 209.
2. Fisher, R. A., *Statistical methods for research workers*, Edinburgh: Oliver and Boyd, 1936. Pp. xiii + 319.

3. KATZ, E., Some factors affecting resumption of interrupted activities by preschool children, *Univ. Minn. Child Welf. Monogr. Ser.*, 1938, **16**, 1–52.
4. LEWIN, K., *A dynamic theory of personality* (Trans. by Adams, D. K., and Zener, K. E.), New York: McGraw-Hill, 1935. Pp. ix + 286.
5. LISSNER, K., Die Entspannung von Bedürfnissen durch Ersatzhandlungen, *Psychol. Forsch.*, 1933, **18**, 218–250.
6. MAHLER, W., Ersatzhandlungen verschiedenen Realitätsgrades, *Psychol. Forsch.*, 1933, **18**, 27–89.
7. OVSIANKINA, M., Die Wiederaufnahme unterbrochener Handlungen, *Psychol. Forsch.*, 1928, **11**, 302–379.
8. SEARS, R. R., Initiation of the repression sequence by experienced failure, *J. exp. Psychol.*, 1937, **20**, 570–580.
9. SNEDECOR, G. W., *Statistical methods*, Ames, Iowa: Collegiate Press, Inc., 1938. Pp. xiii + 388.
10. SLIOSBERG, S., Zur Dynamik des Ersatzes in Spiel- und Ern[illegible]situationen, *Psychol. Forsch.*, 1934, **19**, 122–181.
11. ZEIGARNIK, B., Über das Behalten von erledigten und unerledigten Handlungen, *Psychol. Forsch.*, 1927, **9**, 1–85.

THE CONSISTENCY OF INDIVIDUAL DIFFERENCES IN THE PATTERN OF WORK DECREMENT [1]

BY B. R. BUGELSKI

University of Toledo

Work curves recorded by individuals performing in a weight ergograph differ not only in the quantitative amount of work performed, but also in the qualitative shape or pattern of the curve. The patterns of the curves vary widely from individual to individual. The present experiment was undertaken to determine whether individuals are consistent in the patterns of work decrement they produce in the weight ergograph when the shapes of the curves of work are analyzed in as many ways as possible, both subjectively and objectively. The question as to whether individuals repeatedly produce the same shape of work curve was left in an unsettled state by recent findings of Marks (**4**) which contradict the earlier conclusions of Mosso (**5**) and Yochelson (**8**).

Mosso originally stated that patterns of decrement as produced in the ergograph are consistent for a given individual but that different individuals produce different patterns of work decrement. These statements were based simply on subjective impressions as to the similarity of the shapes of work curves produced by the same individuals. More recently, Yochelson (**9**) came to the conclusion, on the basis of comparisons of curves made by seven subjects, that there is no "typical curve for all subjects, but, rather, a characteristic form for each individual." Yochelson presented sample types of curves (Robinson (**6**), p. 581) with the caution: "These curves, characteristic for each subject, did not set in until late in the practice series, but, after they did appear, were fairly constant from day to day. This observation is in accord with that of Mosso. . . ." No analysis of the different work curves was attempted beyond the admittedly subjective procedure of verbal description by the investigator.

Marks criticized the statements of Mosso and Yochelson because of the subjectivity in ascertaining the degree of similarity between successive work curves of a given subject. Marks further pointed out that but few subjects had been used in earlier studies of ergography, and that the results from these practiced subjects were in-

[1] This is the first of a series of three papers dealing with factors responsible for individual differences in the pattern of work decrement in ergographic work. The three articles were presented in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Yale University in 1938. This experiment was carried out in the laboratories of the Department of Psychology, Yale University. The writer is grateful to the staff of the Department of Physical Education for assistance in obtaining subjects and providing facilities.

sufficient for the broad generalizations that had been made concerning similarity in shapes or work curves. Both criticisms are sound. Marks attempted a more objective attack on the problem of individual consistency in patterns of work decrement. His 164 male subjects, ranging in age from 16 years to 32 years, and in weight from 120 lbs. to 214 lbs., worked with a spring ergograph on two occasions (consecutive days). The subjects were instructed: "Pull back as far as you can each time and make as many pulls as you can." Each subject worked for a specified time (2′ 35″) instead of to complete decrement, because the spring ergograph does not produce complete decrement in the sense that the weight ergograph does. The following measures of work were computed: work rate, change in work rate, contraction height, change of contraction height, irregularity, and tension. The self-correlations between the two separate performances on each of these measures were high, ranging from 0.55 to 0.79. Marks' comment on these correlations is: "This consistency of performance relates not only to those features of the work curve which the subject might be expected to contral voluntarily, *i.e.*, the output measures—work and contraction height—but also to such aspects as irregularity of work and tension which are not measures of output. It would seem that the work curve can yield reliable information about the individual in addition to information about his mere output capacity."

Besides computing the above measures, Marks fitted third order regression equations to the heights of the middle strokes of each 10 second interval. Fisher's (**2**) method of orthogonal polynomials was followed. From these equations numerical constants for the two curves of each of 50 unselected subjects were obtained. Comparison, in terms of correlation of the constants, of the curves for each individual revealed that "the shape of the curve is not constant for a given individual." Marks does not attempt a broad general conclusion, but says: "It is likely that the reported constancy of shape for the weight ergograph actually obtains, and that the present finding represents characteristics specific to the spring ergograph. There exists the possibility, however, that such individual consistency as has been reported for the shape of the weight ergograph may be a product of sampling, of practice, and of subjectivity in ascertaining constancy." Neither of the two resolutions can be accepted until Marks' procedure is carried out on an equal number of unselected subjects with a weight ergograph.

In a more recent study of individual differences in the pattern of work decrement, Yochelson (**9**) attempted to objectify his results by employing judges, unfamiliar with work curves, to sort a large number of curves produced by 111 psychotic subjects. Each subject was

represented by two curves, one obtained with an arm ergograph, and one obtained with a finger ergograph. The curves were matched by the judges for similarity with seven hand-drawn samples. In 54 percent of the cases, the curves made by the same subject on the different ergographs were placed together. Chance selection would have resulted in about 14 percent correct matches. Sample curves produced by five normal subjects with arm, hand, finger, and leg ergographs were reproduced. The subjective similarity of these curves is striking. On the basis of these findings, Yochelson reaffirmed his earlier conclusions (**8**) in which the consistency of individuals in their patterns of work decrement was emphasized.

Marks and Yochelson arrived at contradictory conclusions on the basis of different methods of obtaining and treating their data. The present investigation of the consistency of individuals in the pattern of work decrement is an attempt to resolve the contradictions by employing the techniques of both investigators in the analysis of the data. In obtaining the data the procedures of Yochelson and those who report positive findings for consistency of individuals in ergographic work was followed. If the method of analysis employed by Marks corroborates the conclusions of previous students of weight ergography, an answer to the contradictory findings of Marks will be found in his departure from the more rigorous methods followed by Mosso and Yochelson in obtaining their data.

Apparatus and Procedure

In order to obtain results comparable to previous work on related problems, a modified Mosso finger ergograph was used to secure the work records. Mosso's general scheme (**5**, pp. 83–88) was followed in the construction of the arm rest of the ergograph. The basic construction was of wood and brass. A board 24″ × 9″ was mounted at each end on a triangular wooden block which was so pivoted as to allow the arm rest to be set at an angle of 30° to the right or left. Pitching the arm rest at an angle permitted the subject to work with his arm in a natural position with the dorsal surface of the arm resting on a curved block which was padded with sponge rubber. Movements of the arm were prevented by padded clamps at the wrist and straps near the elbow.

As in other modifications of the Mosso ergograph, clamps were used to secure the first and third fingers. In order to achieve the greatest comfort, these fingers were inserted into large brass tubes mounted on swivels which could be adjusted to any angular spread and so permitted the free exercise of the working finger. Inside the tubes, the fingers rested on sponge-rubber cushions, and were held tight by curved brass plates screwed down from above. Since the third finger normally works jointly with the second, clamping this finger down flat results in great pain when work is fairly under way. To eliminate this pain, the tube for the third finger was mounted on a small wooden block which tilted the tube up at about 30° from the plane of the hand. With the third finger in this partly-flexed position, the subjects rarely complained of pain.

The finger movement used was a flexion of the third finger at the middle joint. Since the palmar surface was exposed, the nature of the movement could be observed and full-finger flexions, from the knuckle, could be prevented more effectively than in finger ergographs where the palmar side of the hand lies flat against the hand rest. The second finger rested on a sponge cushion. The first phalange was strapped down by a half-inch cotton-web strap which restricted all possible movement to the second and third phalanges. A brass ring was used to attach the wire supporting the weights to the finger. The ring was so constructed as to furnish a broad surface (1 cm)

at the point of contact with the finger. The ring was securely taped to the finger at the junction of the second and third phalanges.

The metal track and carriage to which the weights were attached were stock apparatus supplied by Stoelting. A series of weights were used which provided loads up to five kilograms.

It is believed that the precautions taken to ensure the isolation of the movement to the single joint were sufficiently successful to warrant comparability of results, both with previous work and among the subjects here studied. Yochelson (11) has insisted upon rigidity of controls in ergography and his results are in marked contrast to the work of those, notably Marks, who deliberately encouraged lack of control on the assumption that individual differences would be enhanced if a wide range of variability were permitted in the physical aspects of the work.

Recording.—The work curves were traced directly on wax-coated stylograph paper by means of a pointed lever attached to the carriage of the ergograph. Throughout the experiments, a cardboard screen concealed the work record from the subjects.

Timing.—A metronome, set at 60 beats per minute, was placed on the work table in full view of the subject. Contact points on a cross-piece at the base of the metronome completed and broke the circuit from a storage battery to a mercury cup, and thence to an electromagnetic marker which recorded one-second intervals in a line just above the work curve.

Subjects.—Ninety-seven Yale undergraduates who were competing for positions on the crew, track, cross-country, and swimming teams served as subjects. The subjects, ranging in age from 18–22, in weight from 140–210 lbs., formed a fairly uniform group.

Procedure.—Upon entrance into the experimental room the subject was told that a comparison was being attempted of the amounts of work done by athletes and non-athletes and that he was expected to do his best for the athlete group.

The procedure for each subject consisted of the following steps: The subject was seated at the ergograph and adjusted to the apparatus. The clamps and straps were tightened and the subject was asked to indicate points of discomfort if any were present. The following instructions were then read to him: "Bend your finger back at the middle joint as far as you can in time with the metronome. Relax your finger on every other beat. Keep this up until you can no longer bend the joint." The subject was asked if he understood. If the subject indicated doubt, the instructions were reread. Every subject tried a sample lift, and faults, if present, were corrected. The metronome and motor were then started and the subject was told to begin as soon as he had caught the rhythm. The experimenter watched the record until complete decrement had taken place. A ten minute rest was then allowed during which the subject was at liberty to recline on padded rubbing tables in the experimental room. A second work period followed the rest under the same conditions of instruction. All subjects worked with three-kilogram weights on both occasions. Some subjects became interested in the apparatus and were allowed to observe it during the rest period, but no opportunity was given to inspect the work record until after the second work session if such inspection was desired.

Results and Discussion

I. Consistency of Individual Differences in Pattern of Decrement under Standard Conditions

The experimenter carefully inspected the work curves to determine whether the pairs made by each subject corresponded in shape when considered only on the basis of subjective impression. The result of these inspections was in fair agreement with the early contentions of Mosso. The majority of the subjects produced pairs of curves which had the same general characteristics, although the subjects differed widely among themselves.[2]

Because of the fallibility of subjective impressions it was planned to apply objective methods, where the reliability of the results could

[2] Subsequent papers will discuss some of the factors responsible for differences between subjects and also for differences between curves made by the same subject.

be measured, to discover the degree of consistency in the shapes of the pairs of curves made by each subject. Three techniques for determining the degree of consistency were employed: (1) Judgments of similarity to sample types; (2) Analysis in terms of equations of the curves; and (3) Correlations between points in the curves at which contraction height fell below certain values.

1. *Consistency as Determined by Judgments of Shape*

Examination of the 194 curves revealed that they seemed to fall into six different types. These types corresponded fairly well with those reported by Lehmann (**3**), Skawran (**7**), and Yochelson (**9**). Because no one sample curve could be drawn to illustrate the variations within the type, and because individuals who were asked to look at and describe the curves tended to emphasize certain features of the curves at the expense of others, it was decided to objectify the apparent types in terms of three criteria. The judges were thus forced to inspect each curve in terms of the three criteria and come to a decision on the basis of these factors rather than on the basis of some individual undefined criteria. The following criteria were selected: *A*. Duration of the original high level; *B*. Nature of the drop or break in this high level; *C*. Presence or absence of continued work at lower levels.

Various combinations of the above criteria resulted in the following six general types of decrement patterns: 1. Gradual decline, concave from above; 2. Gradual decline, either convex from above, or straight; 3. High level, followed by a break, and a gradual drop from the break; 4. Step-wise curve with a long high level and a shorter low level; 5. Step-wise curve with a short high level and a long low level; 6. High level with a sharp, abrupt decline. Samples of the six types of work curves are shown in Fig. 1*a* and Fig. 1*b*. The characteristics of the six types of curves in terms of the three criteria are given in detail below, in the form in which they were presented to the judges.

I. Gradual decline, concave:
   - *A*. High level not maintained: decrement sets in almost immediately.
   - *B*. No sharp angular breaks.
   - *C*. Decline is negatively accelerated (concave). It may stretch out into the appearance of a tail. An occasional rise in the tail may result in an S shaped appearance.

II. Gradual decline—convex or straight:
   - *A*. Original high level may be maintained for some time or drop gradually.
   - *B*. No definite or sharp break, but a gradual decrement, either positively accelerated (convex), or a continuation of the original rate of decline (straight line).
   - *C*. No level period or tail with the occasional exception of a few short strokes.

III. High level, break, gradual drop:
   - *A*. Original high level maintained for about one-half the curve with the possibility of a drop of a few degrees in this level.
   - *B*. Sudden change in the curve, either as a distinct drop or as a sudden angular change in the slope of the line.

C. Gradual but persistent drop to the base line: in no sense a level period.

IV. Step-wise, *A.*—A two level curve:
- *A.* High level maintained for over one-half of the curve with possibly a gradual drop of a few degrees.
- *B.* Sharp drop to a lower level in a step-like fashion.
- *C.* Lower level continued to the end of the curve with possibly a very gradual decrement.

V. Step-wise. *B.*—Two or three level curve:
- *A.* High level maintained for less than one-half the total length, with possibly a slight drop in this level.
- *B.* Rapid drop to a lower level—usually sharp and distinct, but may lose sharpness by presence of a few strokes of irregular height which may present the appearance of convexity or concavity. More than one step may appear.
- *C.* Lower level continued with gradual or no decrement for the major portion of the curve.

VI. High level, abrupt decline:
- *A.* High level maintained throughout with possibly a slight drop (which may recover) in this level.
- *B.* Abrupt and complete fall.
- *C.* No maintenance of a lower level. Occasionally a few short strokes appear which cannot be considered a continuation of work.

The 194 curves, thoroughly shuffled, were handed to the judge who sorted them into compartments in a wooden framework, each of which was labelled with the description of one of the types of work curves. Each description was accompanied by hand-drawn samples which helped the judges to visualize the features in the verbal description. Nine graduate students in psychology and one undergraduate student served as judges. The following instructions were read to each judge:

> Before you are descriptions of six types of work curves. Each description involves three points which are to be the criteria for judgment. The three points are: *A.* Duration of the original high level; *B.* Nature of the drop from or change in this level; and *C.* Presence or absence of work at lower levels. Study the descriptions carefully and inspect the samples, each one of which illustrates or emphasizes one of the three criteria for each type of curve. Base your judgments on the criteria and not on the samples, which are presented merely to illustrate the criteria. Place each curve before you into that one of the compartments whose description most closely approximates the curve to be judged. Place doubtful curves aside and consider them after the rest have been sorted. Work as rapidly as possible as delay makes for indecision, but remember to analyze each curve according to the three criteria.

Table I presents the number of pairs (two curves made by the same subject) placed into each compartment by each judge, the total number of pairs for all the compartments for each judge, and the chance expectancy of pairs for each judge.[3]

[3] The frequency of correct matches expected from chance factors alone was computed for each judge on the basis of the number of curves assigned to each compartment, by means of the formula:

$$\text{Chance frequency} = \frac{\Sigma Y^2}{(N-1)} \times \tfrac{1}{2} \text{ (Yule, 11)},$$

where Y is the number of curves assigned to any one box by any judge, and N is the total number of curves to be judged, in this case, 194. The critical ratio is obtained by dividing the difference between the actual number of matches and the chance expectancy of matches by the standard error of the chance expectancy—see footnote 5.

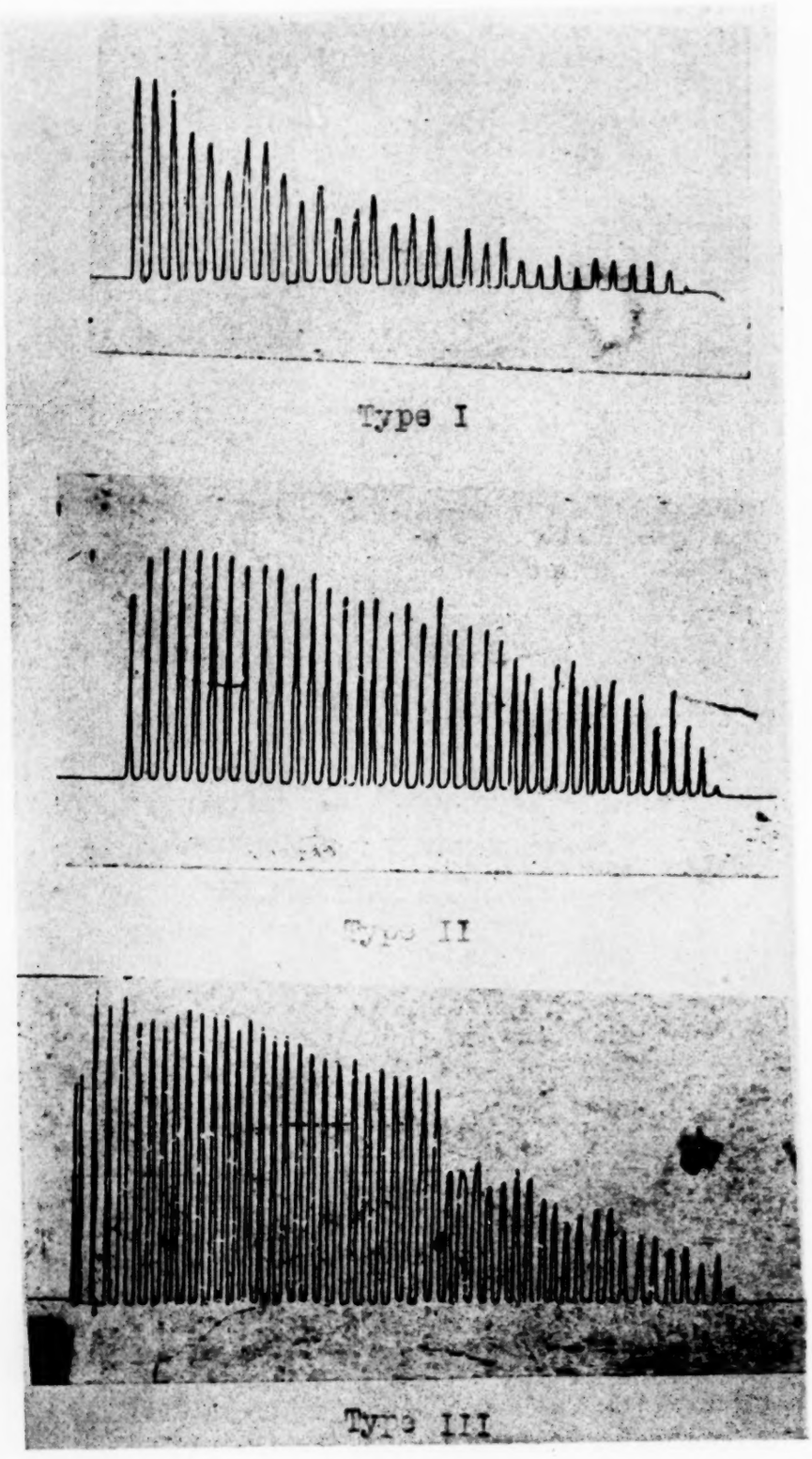

FIG. 1*A*. Ergograph records illustrating the first three types of decrement curves used in the classification by the judges: Type I—gradual decline, concave from above; Type II—gradual decline, convex from above; Type III—high level, break, gradual drop (more detailed descriptions in the text.)

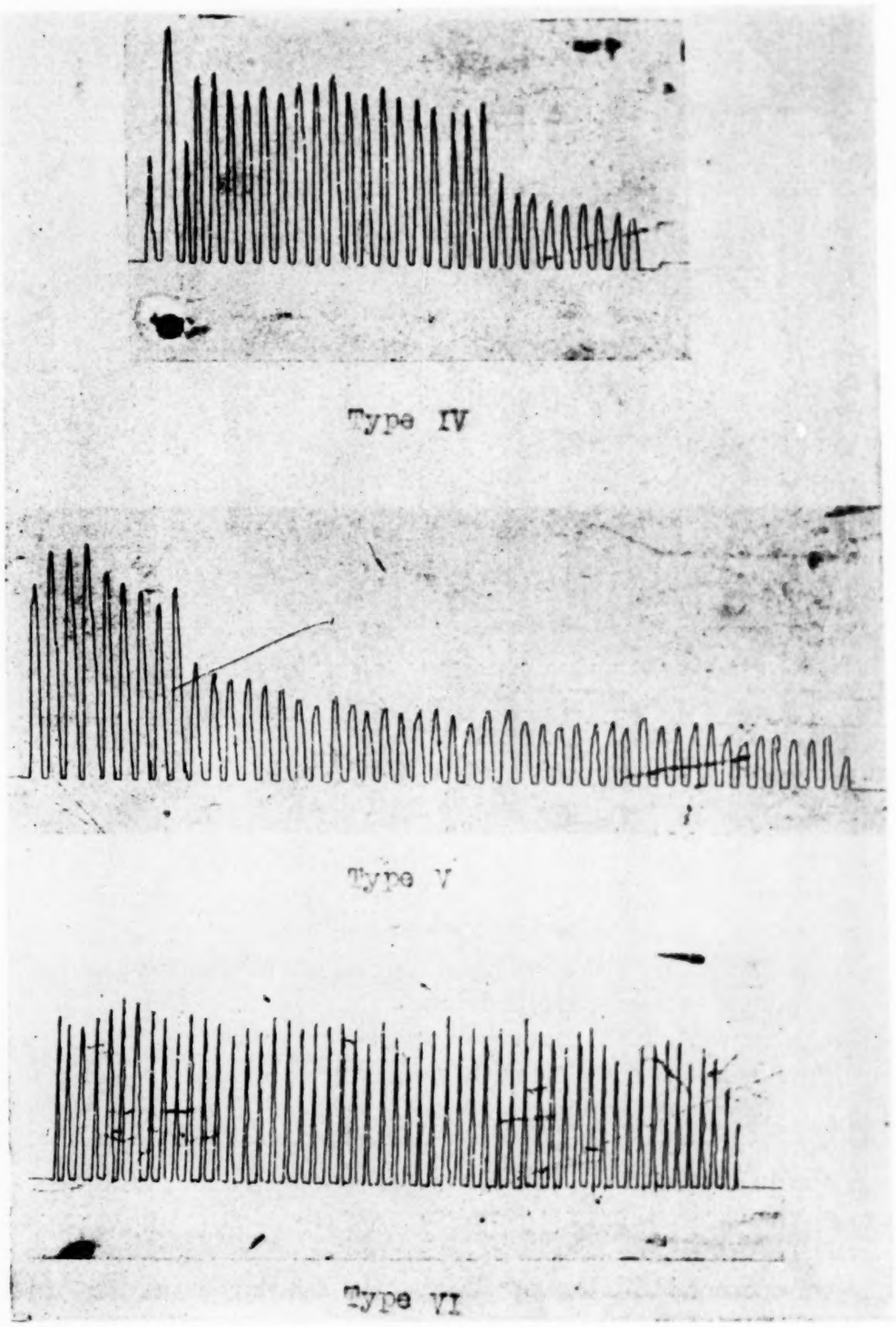

FIG. 1*B*. Ergograph records illustrating the remaining types of decrement curves used in the classification by the judges: Type IV—step-wise *A*, a two level curve (high level over one-half total length); Type V—step-wise *B*, two level curve (high level less than one-half the total length); Type VI—high level, abrupt decline.

TABLE I

THE NUMBER OF PAIRS OF CURVES (CURVES MADE BY THE SAME SUBJECT) PLACED IN EACH CATEGORY BY EACH JUDGE

The total number of pairs for each judge is given in column A; the number of pairs expected by chance is given in column B; the critical ratio of the difference between A and B is given in column C.

Judges	Compartments						A Total Pairs per Judge	B Chance Expect.*	C Critical Ratio *
	I	II	III	IV	V	VI			
A	2	16	6	6	8	8	46	19.83	6.75
B	2	6	7	7	13	8	43	19.50	6.05
C	3	5	2	8	6	7	31	17.31	3.76
D	1	26	7	2	5	6	47	24.84	5.30
E	4	14	2	3	5	9	37	17.86	5.07
F	6	3	0	14	11	12	46	19.83	6.75
G	4	7	7	4	7	8	37	17.06	5.28
H	2	5	9	7	10	7	40	18.09	5.67
I	3	2	7	9	16	7	44	18.78	6.53
J	1	4	0	12	16	5	38	24.01	2.92
Totals	28	88	47	72	97	77	409	196.70	
Means							40.90	19.67	

* See footnote 3.

On the average each judge placed 41 of the possible 97 pairs of curves in the same category. The proportion of correct [4] judgments is 2.07 times its chance expectancy (a perfect score would have been 4.9 times the chance value), and is highly reliable since nine out of the 10 judges showed critical ratios above 3.0 for the difference between their proportion of correct matches and the chance proportion.[5]

The judges agreed that the types defined were represented by the work curves. Each judge experienced hesitation in judgment with about a dozen curves but none of the doubtful curves were of any particular type which would justify the addition of a new category for judgment. Although the types do not in any sense represent a single continuum, there were frequent transitional types or cases which were assigned to one or two compartments quite arbitrarily. The presence of the transitional cases tended to limit the number of correct judgments. The fact that a large number of pairs is assigned to each compartment justified the selection of the type categories.

[4] By correct judgments is meant the placing of two curves made by one subject into the same compartment.

[5] The standard error of the frequency of correct matches expected by chance was determined from the formula: $S.E.^2 = npq$ (Yule 11, p. 257).

2. Results from Mathematical Analysis

Marks' application of Fisher's method of curve fitting by orthogonal polynomials (**2**, pp. 183–186) was followed for the present data.[6] Third-order regression equations were fitted to the heights of the median stroke in each 15th part of 100 work curves produced by 50 subjects selected at random.[7] It will be recalled that Marks' subjects worked for a uniform time (2 minutes, 35 seconds); his curves were divided into 15 periods of 10 seconds each (the first 5 seconds of work were ignored). In the case of the present study, the subjects worked to apparent exhaustion and the lengths of the curves differed. To obtain the desired 15 points of the curve it was necessary to divide the curves into 15 equal parts. Where the number of strokes was not exactly divisible by 15, the remainder of the strokes were distributed uniformly throughout the curve.

The equations obtained were of the form $Y = a + bx + cx^2 + dx^3$, where Y = contraction height and x = time. Each equation contains four coefficients which define the third-order curve of best fit obtainable by this method. The respective coefficients for the two curves produced by each subject were correlated and the reliability of the correlations found. The correlations and indices of reliability ($z/\text{S.D.}_z$) [8] are reproduced in Table II. The same data taken from

TABLE II

SELF-CORRELATIONS OF THE COEFFICIENTS OF THE REGRESSION CURVE OF CONTRACTION HEIGHT ON TIME

Coefficient	r	z	$z/\text{S.D.}_z$
Marks			
A	.778	1.04	7.13
B	.360	.38	2.47
C	.199	.20	1.37
D	.113	.11	.75
Present study			
A	.784	1.06	7.27
B	.685	.84	5.76
C	.425	.45	3.08
D	.288	.29	1.99

Marks is presented with the findings of the present study for purposes of comparison.

[6] The writer is indebted to Dr. Donald C. Perkins for assistance in the mathematical analysis of the work curves.

[7] Actually the curves were selected from two groups of athletes, 50 curves from each group. Since one group comprised only 25 members all the curves from these subjects were used, with no selection. The other subjects were alphabetically chosen from the 40 subjects who comprised the second group.

[8] Z-values were obtained from Fisher (**2**, Table V*B*). When $z/\text{S.D.}_z$ is 3.0 or more, the correlations are regarded as reliably greater than zero.

On the basis of the correlations he found, Marks concluded that there was no constancy for the shape of the curve of decrement in work with the spring ergograph. It will be noted that only one of Marks' correlations is significant, that for the *A* constant which represents, in terms of the work performance, the average contraction height of the strokes at different points in the curve. This coefficient is not particularly meaningful as regards the actual shape of the curve, as the shapes might differ radically without radical changes in average contraction height. The remaining coefficients represent corrections in the equation in order to provide a better fit. In contrast, the correlations for the present study are all higher than those of Marks, and all but the last are significant.

Before any conclusions can be drawn from the reported correlations the adequacy or applicability of this technique must be determined. The only method devised to test the applicability of regression equations to shapes of work curves was that of comparing the correlations for those curves which looked similar to the experimenter with those which did not appear to be similar. The experimenter went over each of the 50 pairs of curves in terms of the criteria established for the judges. The 25 pairs which agreed most closely on the basis of these criteria were selected for comparison with the 25 pairs which agree less closely.[9] The coefficients for the equations of the curves were then correlated as above for the two groups. These correlations are presented in Table III.

TABLE III

SELF-CORRELATIONS OF COEFFICIENTS IN THE REGRESSION CURVE OF CONTRACTION HEIGHT ON TIME FOR THE 25 SUBJECTIVELY MOST SIMILAR PAIRS AND THE SUBJECTIVELY LEAST SIMILAR PAIRS

	Coefficient	r	z	$z/S.D._z$
Most similar N 25.....	A	.773	1.03	4.83
	B	.671	.82	3.84
	C	.588	.77	3.61
	D	.627	.74	3.47
Least similar N 25.....	A	.812	1.13	5.29
	B	.618	.72	3.38
	C	.360	.48	2.25
	D	.251	.26	1.22

The results presented in Table III show that all of the coefficients correlate reliably for the subjectively similar sets of curves, whereas only the first two of the coefficients for the least similar curves correlate reliably. The differences in size and reliability between the

[9] It was more difficult to find curves which lacked subjective impression of similarity than curves which appeared subjectively similar.

two sets of correlations support the validity of the equation technique, and at the same time, validate the subjective impression of similarity. The fact that the *A* coefficient for the less similar sets of curves correlates more highly than the *A* coefficient for the more similar sets does not contradict this conclusion, since the *A* coefficient expresses only the average contraction height.

The results of the mathematical analysis of the shapes of the curves of decrement lead to the general conclusion that individuals are consistent in their patterns of work in weight ergography. Three explanations might be suggested for Marks' failure to secure comparable results; (1) Marks deliberately provided for as much variability as possible (limitation of work to one muscle group was not attempted) on the assumption that individual differences would be magnified and therefore appear more easily. Because of the lack of rigid conditions he may have provided for more variability than he expected; that is, opportunity was provided for individuals to vary their performance the second time they worked. (2) The results reported by Marks may hold for the spring ergograph, which because of its nature does not produce complete exhaustion but invites work on a lower level of performance for most subjects and therefore tends to call forth a typical curve for all subjects. If all the subjects in Marks' esperiment worked fairly uniformly and showed no great differences in type a high correlation between the constants in the equations would not be expected. (3) Marks selected the middle stroke of each interval as a 'random sample' for that interval. It is possible that further variability was introduced by this method. In the present study the median stroke of each interval was used. It is unfortunate that Marks avoided any subjective analysis of his curves, as such an analysis might have provided the solution to the contradictions between his results and previous studies.

3. *Results from Self-correlations of Points on the Work Curves at Which the Decrement Reached Certain Values*

Subjective inspection of the work curves revealed that very often sharp and abrupt drops occur at about the same relative positions in the two curves made by a given individual. This might find quantitative expression in the correlation between the points in the two work curves made by each individual at which the level of performance fell below a certain value. Two values were chosen, 75 percent and 50 percent of the initial stroke height, for study in this way. No meaningful correlations could be obtained by using percentages much below 50 percent because of the lack of scatter in the lower ends of the curves. The same curves which were used in the mathematical analysis reported above were used because the medians

for each 15th of the curves were available. The coefficient of correlation was computed between the points in the two curves produced by each indidivual in which the level of stroke height fell below 75 percent of the initial value. A similar correlation coefficient was computed for the 50 percent value. The correlations are presented in Table IV.

TABLE IV

THE SELF-CORRELATION OF POINTS IN THE TWO CURVES MADE BY EACH SUBJECT AT WHICH THE LEVEL OF THE STROKE HEIGHT FELL BELOW: (1) 75 PERCENT OF THE INITIAL VALUE; (2) 50 PERCENT OF THE INITIAL VALUE

No. of Pairs	Level of Stroke Height	r	P.E.$_r$
50	75% of initial value	0.50	0.07
50	50% of initial value	0.60	0.06

Both correlation coefficients are significant and are high enough to confirm the findings on judgments and mathematical curve analysis. The correlations indicate a tendency for a given subject to reach the same level of decrement at the same time on two different occasions of work. It has been the experience of the present investigator, as well as of others who have worked with the same subjects on several occasions, to expect or anticipate points in the performance at which radical drops would occur. Such subjective anticipations are borne out frequently enough to form some basis for the conclusions of early investigators as to the essential similarity of successive performance on the ergograph. The correlations reported above confirm this subjective observation.

SUMMARY AND CONCLUSIONS

Ninety-seven male athletes worked on two separate occasions with a modified Mosso ergograph. The patterns of work decrement in the resulting work curves were analyzed in three different ways: (1) The curves were sorted by 10 judges according to objectively defined criteria. (2) Third-order regression equations were fitted to the heights of the median strokes in each 15th of the pairs of work curves for a random sample of 50 subjects; evidence is reported for the validity of this procedure. (3) Points along the work curves at which the level of performance fell to 75 percent and 50 percent of the initial stroke were computed.

The results of the judging technique demonstrated a highly reliable degree of consistency between the patterns of decrement in the two work curves of the individual subjects. Self-correlations between the coefficients of the regression equations were high and reliable,

indicating the same conclusion as do the results of the judging. Self-correlations between points at which the level of work was 75 percent and 50 percent of the initial value were also significant (above 0.50); these correlations further confirmed the conclusion that individuals are consistent in their manner of arriving at complete decrement in ergographic work.

(Manuscript received September 27, 1940)

References

1. Crawley, S. L., An experimental investigation of recovery from work, *Arch. Psychol.*, 1926, **13**, No. 85, pp. 66.
2. Fisher, R. A. *Statistical methods for research workers.* London, Oliver and Boyd, 1934, 5th ed., pp. 319.
3. Lehmann, H., Messung des Kraftimpulses, *Indus. Psychotechn.*, 1928, **5**, 264–272.
4. Marks, E. S., Individual differences in work curves, *Arch. Psychol.*, 1935, **27**, No. 186, pp. 60.
5. Mosso, A. *Fatigue.* (Trans. by M. Drummond and W. B. Drummond.) New York, Putnam, 1904, pp. xiv + 334.
6. Robinson, E. S. Work of the integrated organism, *Handbook Gen. Exper. Psychol.*, C. Murchison ed., Worcester, Mass., Clark Univ. Press, 1934, 571–650.
7. Skawran, P., Die Typologie der Ergogramme und ihre Beziehung zu Energietypen und zu den Korperbautypen Kretschmers, *Psychotechn. Z.*, 1931, **6**, 10–19.
8. Yochelson, S. Effects of rest-pauses on work decrement. Yale Univ., Ph.D. dissertation, May 1, 1930, pp. 102.
9. ——, Individuality in ergography, *Amer. J. Psychol.* (In press).
10. ——, Mechanical controls of arm, leg, and hand ergography, hand dynamometry, *Amer. J. Psychol.* (In press).
11. Yule, G. U. *An introduction to the theory of statistics.* London, Griffin, 1932, pp. 434.

AN EXPERIMENTAL ANALYSIS OF THE LAW OF EFFECT [1]

BY HANS WALLACH AND MARY HENLE

Swarthmore College

Thorndike has presented a large quantity of evidence "to show that a satisfying after-effect of a connection does in fact strengthen it under conditions equalized in respect of all other forces than the satisfying after-effect." [2] The experiments reported here were performed to determine whether or not it is the satisfying after-effect *alone*, the success of the reaction itself which 'strengthens connections' and thereby brings about repetition of the successful reaction.

The Law of Effect has been tested with human subjects only in situations in which it was the obvious task of the subject to learn the rewarded material. Repetition of rewarded responses has so far been shown to occur only where such repetition was sensible; it has not been demonstrated merely as an automatic effect of reward.

The experiments to be reported in the present paper were designed to present a situation in which repetition of rewarded responses was possible, while no reason for or against such repetition existed for the subject. Only if under these conditions rewarded responses are still repeated more frequently than by chance, such repetition may be attributed exclusively to an effect of the reward.

Such a situation was created by adopting with modifications a procedure of Thorndike's. He had prepared a list consisting of pairs of words and numbers. The numbers were chosen from a limited sequence of numbers and arbitrarily assigned to the words. A word was read to the subject, who was to guess the corresponding number from this sequence. He was immediately told whether his response was correct or incorrect, and the experimenter proceeded to read the next word of the list. When the list was thus completed, the experimenter started again with the first word. Sooner or later in the course of the repeated presentations of the list the subject was bound to make some hits by chance. Once a correct response was found, the subject was to repeat it on subsequent presentations of the list. Correct responses were found to be repeated from presentation to presentation to an extent far exceeding chance expectation. Similar results have been obtained in many previous experiments.

[1] The writers wish to express their gratitude to Professor W. Köhler for his criticism of this paper.

[2] Thorndike, E. L. An experimental study of rewards. *Teach. Coll. Contrib. Educ.*, No. 580. New York: Teachers College Bureau of Publications, 1933. P. 2.

This result was not, however, Thorndike's chief concern in the experiment under discussion. In addition to the repetition of correct responses, Thorndike found that incorrect responses were repeated more frequently than they would have been repeated by chance, and the more so the closer they were in the series to a correct response. It was this gradient of repetition of wrong responses adjacent to a correct one to which Thorndike called attention. He considers it significant because he believes that it "proves that a satisfier or reward can act directly and immediately to strengthen a connection."[3] While 'inner repetitions' or 'expectations' on the part of the subject may complicate the direct action of a satisfier on a rewarded connection, Thorndike believes that in the case of the strengthening of a neighboring punished connnection such complications are absent, and "the satisfier acts upon it unconsciously and directly, much as the sunlight acts upon plants, or an electric current upon a neighboring current, or the earth upon the moon. Presumably it acts upon the connection which is psychologically nearest to it of all, the connection whose after-effect it is felt to be, primarily in this same immediate and direct manner. A satisfying after-effect can strengthen its connection with no paraphernalia to help it. . . ." [4]

This type of experiment, since it is Thorndike's best case for the automatic action of rewards, was chosen for the present analysis. Our experiment duplicated Thorndike's in all important respects except the following: The instructions were framed so that no sensible reason for the repetition of correct responses was given to the subject. Only the satisfying after-effect of these responses could then be responsible if successful reactions should be repeated more frequently than by chance.

Such an experiment, in which the learning motive is to be excluded, and where, therefore, the conscious activity of the subject must be restricted to mere guessing, obviously needs *some* motive. Otherwise the nature of the psychological situation would become entirely undefined. However, this motive may by no means interfere with the purpose of the experiment, which can test a direct influence of the after-effect only if the subject is left entirely free to repeat or not to repeat a given response. These requirements were met by presenting the experiment to the subject as a test of extra-sensory perception.[5]

[3] Thorndike, *op. cit.*, p. 47.

[4] *Ibid.*, p. 48.

[5] No assumption about the existence of extra-sensory perception is made here. This experiment has the character of a test of extra-sensory perception only for the subject. It will appear below from a detailed description of our procedure that actually the results of our experiment could not have been influenced by such an ability should it exist.

This type of procedure appears to be particularly suitable for our purposes also for another reason. If, in a test of extra-sensory perception, the subject tends to assume a conscious attitude of being receptive to the presumably unnoticeable influences of extra-sensory perception, such an attitude should also favor the operation of an automatic effect of reward on the subject's guessing.

The following instructions were read to the subject: "You have perhaps heard that experiments on extra-sensory perception are being performed by Mr. Rhine at Duke University. Subjects are asked to name items which they cannot possibly perceive in any ordinary sense of the word. If they are able to make hits more frequently than they would by chance alone, it must be by extra-sensory perception.

"We are dealing with the same problem by a method somewhat different from Mr. Rhine's. We have prepared a list of pairs of words and numbers like this." [Here a sample list of word and number combinations was shown to the subject.] "We shall read you the words, and you guess the numbers that go with them. If extra sensory perception exists, you should be able to do this better than chance. If it does not exist, you will merely be guessing the number that goes with each word, and you should not make a score better than chance.

"As numbers we have used the digits from 1 to 10. So you can choose among ten different reactions in each case, and with this large number you will, of course, not make many hits. But neither did Rhine's subjects. After each judgment we shall tell you whether you were right or wrong. To encourage you further, we shall pay you half a cent for each hit you make.

"Thinking about your performance during the experiment will be definitely detrimental in a function as intangible as extra-sensory perception. So try to behave naively and give your responses as quickly as you can. Do not mind if we hurry you through the experiment; it will only help you to make a better score.

"We have a list of 20 words, and this list will be presented a number of times. But the numbers assigned to the words will change in a random fashion from presentation to presentation. Thus a number which is correct on the first presentation will not necessarily be correct on the next."

Two experimenters conducted the experiment, the one reading the words to the subject and informing him of the correctness of his responses, the other recording the responses. The general speed of procedure, as indicated by the average time between successive responses, was comparable to that employed by Thorndike. As subjects, 50 undergraduate or graduate students were used.

We may summarize our procedure as follows: The present experiment duplicated Thorndike's in all important respects except that a learning motive was not present, as in Thorndike's experiment; the experiment was rather, for the subject, a test of extra-sensory perception, in which there existed, therefore, no reason for him to repeat previous correct responses.

The experiment was designed to answer in particular the following questions: How great is the number of repetitions of rewarded responses here as compared with that which Thorndike found in his experiments? How does it differ from the number of repetitions of wrong responses? Does a gradient of repetition of wrong responses around the correct ones exist?

To facilitate computation of such a gradient, we borrowed a procedure first used by Muenzinger and Dove.[6] Instead of using a prearranged list of correct word-number combinations, which would produce a random distribution of hits, we used a prearranged pattern of positions of the responses which were to be called correct no matter

[6] Muenzinger, K. F. and Dove, C. C. Serial learning: I. Gradients of uniformity and variability produced by success and failure of single responses. *J. gen. Psychol.*, 1937, **16**, 403–413.

what the subject's reaction was. Thus we were able to get evenly spaced correct responses. Each presentation of the list contained three rewarded responses; any two of these were separated from each other by six incorrect ones. This gave us a list of twenty-one words. The positions of the correct responses were varied in an orderly manner on consecutive presentations of the series. Four series containing no hits were inserted. In all, the list was repeated twenty times. Ten subjects were used in Experiment I.

The results of Experiment I are presented in Table I. Rewarded responses were repeated on the presentation immediately following the reward in 10.4 percent of the cases. This result is to be compared with a percentage of repetition of correct responses of 'about 50' in Thorndike's experiment.[7] Wrong responses were repeated in 10.8 percent of the cases under our conditions. The difference between 10.4 percent and 10.8 percent is completely insignificant. Experiment I thus shows no difference of frequency of repetition between rewarded and wrong responses.

TABLE I

REPETITION OF REWARDED AND WRONG RESPONSES

	Rewarded Responses			Wrong Responses		
	No. of Responses	No. of Repetitions	Percent of Repetitions	No. of Responses	No. of Repetitions	Percent of Repetitions
Expt. I.....	472	49	10.4	3706	402	10.8
Expt. II.....	823	95	11.5	5262	656	12.5
Total.....	1295	144	11.1	8968	1058	11.8

Data obtained under exactly comparable conditions in another experiment, which will be described below, are pertinent here, and may be added to the results of Experiment I.[8] Again the difference between rewarded and wrong responses is negligible. The combined results show that rewarded responses were repeated in 11.1 percent of the cases, wrong responses in 11.8 percent. *Rewarded responses were thus not repeated more frequently than wrong ones.*

[7] Thorndike, *op. cit.*, p. 4. It is apparent from the description of the procedure we employed that our experiments are not strictly comparable with the experiment of Thorndike's which has been described above, but with an otherwise similar experiment in which the positions of the correct responses are rotated from trial to trial. Such an experiment has been performed by Muenzinger and Dove, but the results are not numerically comparable to ours since their list of words was twice as long as ours, and should, therefore, give a lower percentage of repetition of correct responses. However, their experiment, as Thorndike's, showed repetitions of correct responses to be high above the level of chance expectation. *Cf.* Muenzinger and Dove, *op. cit.*

[8] The additional data came from experiments with forty subjects. Fewer presentations of the list were given here than in Experiment I.

It will be noted that the percentages of repetition of wrong as well as of rewarded responses (11.8 and 11.1 respectively) are slightly above those which chance alone would have given (10.0). A possible explanation is to be found in the fact that subjects restricted the range of numbers from which they selected their responses to a range smaller than the ten on which calculations of chance probabilities were based. If theoretical probabilities had been based on the range of numbers actually used by the subjects, it is likely that the percentages obtained would not differ from chance. In any case, the important comparison is between the percentages of repetition of rewarded as against wrong responses, and here no significant difference exists.

In yet another respect the results of our experiments differ from Thorndike's. Although our experiment was designed specifically to bring out a gradient of repetition of wrong responses adjacent to rewarded ones, no such gradient appears. The relevant results from the two experiments are presented in Table II. The slight differences

TABLE II

REPETITION OF WRONG RESPONSES AT DIFFERENT DISTANCES FROM THE REWARDED ONES

	Rewarded Responses			Wrong Responses One Place Removed			Wrong Responses Two Places Removed			Wrong Responses Three Places Removed		
	No. of Responses	No. of Repetitions	% of Repetitions	No. of Responses	No. of Repetitions	% of Repetitions	No. of Responses	No. of Repetitions	% of Repetitions	No. of Responses	No. of Repetitions	% of Repetitions
Expt. I	472	49	10.4	944	96	10.2	937	93	9.9	934	96	10.3
Expt. II	823	95	11.5	1,646	219	13.3	1,646	205	12.5	1,636	198	12.1
Total	1,295	144	11.1	2,590	315	12.2	2,583	298	11.5	2,570	294	11.4

in repetition of wrong responses at different positions with respect to rewarded responses are not significant. These results are based on numbers of cases which closely approximate those used in the experiments already referred to in which Thorndike as well as Muenzinger and Dove obtained gradients.

Our failure to obtain a gradient is not astonishing in view of the absence of a high level of repetition of rewarded responses in our experiment. Under conditions where the gross effect of reward on the correct response itself is lacking, we can hardly expect to demonstrate its subtle influence on neighboring responses in the form of a gradient. Since, in Thorndike's view, the gradient is a direct consequence of the reward, just as is the repetition of rewarded responses, and since the latter is by far more conspicuous in Thorndike's results, we shall center our discussion of the effect of reward on the case of repetition of rewarded responses.

Also much more conspicuous than the gradient as such is the general high level of repetition of wrong responses in Thorndike's

experiment. There wrong responses five or more steps removed from a rewarded response (thus far enough away not to be subject to the 'strengthening' influence of the reward) were repeated in about 20 percent of the cases; that is twice as frequently as they would have been repeated by chance, and well above the level of repetition of either rewarded or wrong responses in our experiment. Thorndike considers this to be an effect of the mere occurrence of the response. He writes in this connection: "In such experiments the fact that a person responds to a word by a certain number makes him more likely to respond to that word at the next trial by that same number, even though the response was punished. The connection is strengthened more by being made than it is weakened by being punished." [9] We failed, however, to obtain a level of repetition of wrong responses at all comparable to this. This finding seems to exclude the explanation which Thorndike gives for his results. Even if our subjects reacted under the influence of a tendency to vary their responses, a possibility which will be discussed below, it is not plausible to argue that this tendency was stronger than the tendency of Thorndike's subjects to vary their wrong responses.

Thus our results are entirely different from Thorndike's, although the only important respect in which the two procedures differed was that in Thorndike's experiments the subject was led to expect a right response to be correct on subsequent trials, while in our experiments this was not the case. The high percentage of repetition of correct responses and, to a lesser degree, of wrong responses, and the gradient, which Thorndike obtained, must thus be dependent upon this difference. Of the two variables present in Thorndike's experiment, a possible automatic effect of the reward and the learning motive present in the situation, the latter seems to be the decisive one for his results. If it is eliminated, no automatic effect of reward appears to be demonstrable.[10]

[9] Thorndike, *op. cit.*, p. 4.

[10] It has been suggested to us that our failure to obtain a high level of repetition of rewarded responses is a consequence of a lack of 'belonging' between stimulus and response. Thorndike has shown that another condition of learning, at any rate, namely repetition, does not 'strengthen connections' in the absence of belonging. As Thorndike defines belonging, however, it was certainly present in our experiment. He writes: "The belonging which is always or nearly always necessary in order that the repeated occurrence of a sequence may strengthen the connection between the first term of the sequence and the second need not be more than the least which the word implies. There need be nothing logical, or essential, or inherent, or unifying in it. Any 'this goes with that' will suffice. Each nonsense syllable in a series which is read as a series 'belongs' to the one before it in the series. 1492 belongs to Mr. Jones as his telephone number as truly as to Christopher Columbus as an auspicious year. In an experiment, 1492 may truly belong to 65 or 7843, or sig nop." (*The Fundamentals of Learning*. New York: Teachers College Bureau of Publications, 1932. P. 72.)

Thus, in Thorndike's sense, belonging is just as truly present in our experiment as in his. Our results can only be attributed to a lack of belonging if, in a definition of belonging, one should include some reference to the motivation of the subject.

It might be argued, however, that the results of the present experiments can be explained by a possible tendency against repetition of previous responses on the part of the subject. The low level of repetition of correct responses would, according to this view, be a result of two opposing tendencies: an automatic effect of reward and a tendency of the subject to avoid repetition of rewarded responses. Such a tendency, it might be argued, could be created by the instruction: "The numbers assigned to the words will change in a random fashion from presentation to presentation. . . ."

This interpretation can be subjected to an experimental test. The possibility of the operation of such a tendency presupposes that the subject remembers the preceding correct responses in order to avoid them. Experiment II was set up to determine the subject's memory of correct responses apart from his repetition of them. The procedure of Experiment I was followed without change for six complete presentations of the list. Then a test of recall was introduced with the instructions: "Now will you please repeat the responses you just gave while I read the list of words. I mean try to remember your last responses." After this interruption the series was continued as before, in the case of 20 of our subjects, to obtain more data on repetition of rewarded and wrong responses.[11]

Forty subjects were employed in this experiment. For 10 of them, as in the rest of the experiment, 3 out of 21 responses were made correct in the presentation of the list immediately preceding the test of recall. For the other subjects the number of correct responses on this presentation was increased to four in order to obtain more data on correct responses; here the list was reduced to twenty words to secure even spacing of rewarded responses.

The amount of recall exhibited by the subjects in this test varied considerably. Sixteen subjects were found whose level of recall for all the items of the test was not above chance (0 to 2 hits). On the recall test these subjects reproduced only 6 percent of their previous responses, while chance would have given hits in 10 percent of the cases. Recall of previously rewarded responses alone was likewise not above chance for these subjects. Thus they cannot, on the whole, be assumed to show any memory for their previous responses. With respect to repetition of responses in the other part of the experiment ('extra-sensory perception'), these subjects did not differ from the others, and they can be regarded as representative of the whole group of subjects. To argue, now, that in these subjects an automatic tendency to repeat correct responses was offset by a tendency to avoid

[11] These data on spontaneous repetition of responses, it will be recalled, are presented in Tables I and II as the results of Experiment II. There the results from the two halves of the experiment (before the recall test and after) were combined since they showed no difference.

such repetition is impossible. Both an automatic effect of reward and an intentional variation of response imply memory of the items, and the test has shown no actual basis in the memory of these subjects for either effect. It is thus highly unlikely that the low level of repetition of rewarded responses obtained can be explained by the operation of a tendency to avoid such repetition.

The test of recall has, however, another and much more direct significance for the interpretation of our main experimental result. We have as yet only demonstrated that, under the conditions of our experiments, reward does not bring about the repetition of correct responses. It is, however, conceivable that reward has an influence on memory which, for one reason or another, does not become overt in the behavior elicited by our 'extra-sensory perception' procedure. The test of recall of Experiment II was introduced to demonstrate directly a possible effect of reward on memory as revealed in intentional recall, apart from any effect it may have on responses which are made in the absence of an intention to recall. Such an effect would be demonstrated if, under the instruction to recall, rewarded responses were found to be remembered better than wrong ones. A comparison of recall of responses rewarded on the previous trial with that of responses previously called incorrect revealed no significant difference (Table III). Rewarded responses were recalled in 18 per-

TABLE III

RECALL OF REWARDED AND WRONG RESPONSES IN EXPERIMENT II

	Rewarded Responses	Wrong Responses
No. of responses	150	660
No. recalled	27	124
% recalled	18.0	18.8

cent of the cases, wrong ones in 18.8 percent. This result is too high to be regarded merely as a matter of chance hits. If rewarded and wrong responses show no difference in reproduction on the test of recall, this must therefore be considered as a valid finding regarding their state of memory.

Since subjects differed rather widely in their performance on the test of recall, it seemed advisable to deal separately with the results of subjects who exhibited a large amount of recall. The results of fifteen subjects, each of whom made five or more hits in the test of recall, were considered. These subjects made 100 hits out of a possible 304 on the test of recall.[12] Rewarded responses were remembered in 32 percent of the cases, wrong ones in 33 percent. These results again show no difference in recall between rewarded and wrong

[12] *Cf.* above. Some of these subjects were given lists of 21 words, the others lists of 20 words.

responses. Yet, if Thorndike were correct in assuming an automatic effect of reward on memory, it should have expressed itself in terms of a difference in recall of responses which on the preceding presentation in the 'test of extra-sensory perception' were rewarded or punished. Intentional recall, however, as well as response without the intention to recall, has failed to show a selective effect of reward.

With the demonstration that *reward as such has not the effect of favoring recall*, there appears to remain no possibility of attributing Thorndike's results to an automatic effect of reward. When no learning motivation is introduced, repetition of rewarded responses as well as of wrong ones has been shown to remain on a chance level, although reward is present. In Thorndike's experiments the high level of repetition of rewarded responses must therefore depend upon the presence of this learning motivation.[13]

How, then, can we interpret the chief result of Thorndike's experiment, namely the high level of repetition of rewarded responses? It can perhaps in part be attributed to an intention on the part of the subject to *reproduce* at their proper places as many correct responses as possible on successive trials. Such a tendency was doubtless present under Thorndike's conditions, but was, of course, absent in our 'test of extra-sensory perception,' where the subject merely guessed. That such an intention alone can actually raise the level of reproduction of responses is shown by our experiment in which a test of recall was introduced. In that test the number of reproductions was raised from a level of about 12 percent to about 18 percent when the subject was suddenly asked to repeat responses, which had been made, it will be remembered, in the absence of a motive to learn them. Yet it is clear from a comparison of the latter figure with the number of repetitions of correct responses which occurred in Thorndike's experiment (about 50 percent) that an intention to reproduce particular responses cannot alone account for this number. There was also present in Thorndike's experiment, and absent in ours, the motive to *learn* the correct responses in order to repeat them later. An effort on the part of the subject to commit correct responses to memory probably chiefly accounts for the high level of repetition of correct responses in Thorndike's experiment. As must also be the case with the intention to reproduce correct responses, this effort to learn the correct responses results from a motivation which probably prevails throughout the experiment, and is not created individually each time a reward occurs. What Thorndike and others call reward or success, the reaction of the experimenter to the subject's response

[13] It is clear that here the term motivation does not designate a driving force which brings about action within a given situation, but rather the 'meaning' of the situation itself which directs behavior in the situation in a particular way.

('right' or 'wrong'), seems, then, to be merely a sign which indicates to the subject which responses to commit to memory.

A discussion of the significance of these results for animal learning must be postponed until a better appreciation of the psychological consequences of the learning motive in a situation is possible. If, in the absence of a learning motive, reward shows not the slightest effect on an individual stimulus-response connection in human beings, it seems appropriate to view very critically any supposition of an automatic effect of the reward in animal learning.

Summary

Experiments were performed to determine whether repetition of rewarded responses is, as Thorndike holds, an automatic effect of the reward alone, or whether it is a consequence also of the learning motive present in the situation which Thorndike employed. An experiment was devised which duplicated Thorndike's in all important respects except that it did not possess the character of a learning experiment. Under these conditions, results were obtained which differed from Thorndike's in the following respects: (1) They failed to show the high percentage of repetition of rewarded responses which Thorndike found. Rewarded responses were not repeated more frequently than wrong ones. (2) The general high level of repetition of wrong responses which was conspicuous in Thorndike's experiment was likewise absent here. (3) No gradient of repetition of wrong responses around rewarded ones was found. A control experiment showed that these results cannot be attributed to a tendency of the subject to avoid repetition of previously rewarded responses. This experiment likewise eliminated the possibility that the reward did have an influence on memory which, for one reason or another, could not become overt in the behavior elicited by the procedure in our first experiments. The present experiments may thus be interpreted as demonstrating that reward as such does not have the effect of favoring recall.

(Manuscript received October 6, 1940)

EXPERIMENTAL EXTINCTION OF AN HALLUCINATION PRODUCED BY SENSORY CONDITIONING[1]

BY DOUGLAS G. ELLSON

Stanford University

In a previous paper (**6**) an auditory hallucination was produced by a procedure of repeatedly pairing a cue-stimulus (light) with a tone. The question was raised as to whether the hallucinations obtained were conditioned responses in that their relations to other variables were the same as those found in 'typical' conditioned response experiments. The present experiment investigated the effect of two of these variables: (1) the omission of the tone, or 'unconditioned stimulus,' and (2) the subjects' expectations of the tone as determined by instructions.

Thirty subjects, seven male and twenty-three female university undergraduates, took part in the experiment. The apparatus and training procedure are the same as those used in the previous experiment (**6**), so only a brief description of essentials will be presented here. A 1000-cycle tone was used as the unconditioned stimulus (UCS), the tone being varied in intensity by 2-decibel increments. A light on the arm of the subject's chair furnished the conditioned stimulus (CS). A telegraph key on the right arm of the chair was connected with a signal in the experimenter's room.

The 'Call for Subjects' circulated in the class-room described the experiment as a test of the ability to hear faint tones. When the subject came to the experiment, he was told to be seated in the experimental chair. The tone was turned on at an intensity about 20 db above threshold, and the following instructions were given:

> The tone which I have just turned on is the one I am using in this experiment. I want to determine how faint a tone you can hear. What I want you to do is to press that key as soon as you hear the tone, hold it down as long as you hear it, and let it up when you don't hear it any longer. Please pay very close attention because the tone is very faint most of the time. The only other thing you have to do is to keep your eyes open so that you can see that light, which is a ready signal. You don't have to watch it; just be sure you can see it when it comes on.

Sixty training trials were given, with three reinforced test trials after the 20th, 40th, and 60th training trial. In training trials, the light and tone were presented simultaneously, the tone, however, being initially below threshold. It was increased in intensity at the rate of 2 db/sec until the subject pressed the key to indicate that he heard it, then decreased at the same rate until he released the key, when both light and tone were terminated. The intervals between the onset of the light and the subjective onset of the tone were kept roughly between two and six seconds by controlling the initial (sub-threshold) intensity of the tone to compensate for slow fluctuations in the threshold. The reinforced test trials were similar to the training trials except that fifteen seconds elapsed between the onset of the light and the sub-threshold onset of the tone, permitting the subject to respond to any hallucinatory tone heard during the interval. If no response was given, the tone was presented as usual; if the subject pressed the key in the 15-second test interval, the tone was presented so as to blend with the hallucination, then tapered off as in the training trials.

Following the last of the reinforced test trials, a series of ten extinction trials was given in which the light was presented alone for thirty seconds, or, if the subject pressed the key within

[1] This paper has been accepted and approved by the Stanford Committee on Psychical Research as Communication No. 9 from the Psychical Research Laboratory.

the 30-second period, until he released it. If the subject gave no hallucinatory responses in the ten trials, he was given a few presentations of the light and tone together, and dismissed. This procedure was followed in order to prevent the subject from becoming aware of the true nature of the experiment as a result of a long series of trials in which the tone did not occur. If the subject gave one or more hallucinatory responses in the first ten trials, the series was continued with ten more, making a total of twenty extinction trials.

The experimenter then entered the subject's room and gave new instructions approximately as follows:

> This last series of tones has been very weak, and it is possible that you may have been imagining the tone part of the time or pressing the key when some other noise occurred. To check on that I am going to give you a series of trials in which the tone will not be given. I will present the light just as I have been doing until now, but the tone will be omitted. I want you to press the key just as before if you hear the tone or noise which you have been pressing the key for.

These instructions were elaborated if necessary to be sure that the subject understood the expressed purpose of the test and what he was to do. Then ten more 30-second extinction trials were given similar to the preceding twenty trials.

The steps in the procedure may be summarized as follows:

I. Initial instructions.
II. 20 training trials.
III. 3 reinforced test trials (15-second interval).
IV. 20 training trials.
V. 3 reinforced test trials (15-second interval).
VI. 20 training trials.
VII. 3 reinforced test trials (15-second interval).
VIII. 20 extinction trials (30-second). (Subjects not giving one or more hallucinatory responses in the first 10 trials dismissed.)
IX. New instructions. Subjects told that no tone will be given.
X. 10 extinction trials (30-second).

Results

As indicated above, subjects who did not give hallucinatory responses in any of the first ten extinction trials were dismissed. In the remainder of the paper, results are given and discussed for only the remaining subjects, *i.e.*, those who gave responses during extinction trials and completed the experiment.

Table 1 presents the latency of each hallucinatory response given in the first 20 extinction trials, together with frequency and mean latency for each subject and for each trial. Table 2 presents similar data for the duration of response. These data are summarized graphically in the left-hand portion of Fig. 1. Progressive increases in latency or decreases in frequency and duration, such as would presumably be expected in an extinction series, are not present.

Table 3 presents data for latency, duration, and frequency during the ten trials occurring after the subjects had been told that no tone would occur.[2] The grand means for this table, obtained by averaging the subject means, are compared in Table 4 with the corresponding means obtained from the data in Tables 1 and 2. This comparison

[2] The last column in Table 3 (Answers to Questionnaire) will be treated in a later section.

TABLE 1

LATENCY IN SECONDS OF HALLUCINATORY RESPONSES ON 20 SUCCESSIVE EXTINCTION TRIALS

The absence of response is indicated by a dash.

Subject No.	Sex	Trial No.																				Total Frequency	Mean Latency
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20		
1	M	—	—	21	15	5	11	8	5	2	21	8	6	5	11	6	6	3	7	5	13	18	8.78
3	M	—	—	—	—	15	5	29	—	—	—	2	—	—	—	6	26	—	—	—	22	7	15.00
6	F	—	—	—	—	—	—	—	—	7	8	—	13	5	20	19	25	—	17	—	15	9	14.33
10	F	—	—	—	—	—	16	—	—	—	—	—	—	—	—	9	4	19	26	—	—	5	14.80
11	M	—	—	18	2	18	—	11	11	26	6	8	6	25	—	21	—	—	—	22	—	12	14.50
12	F	7	—	11	—	—	6	3	14	8	—	8	13	12	—	—	—	19	2	—	5	12	9.00
13	F	4	2	3	3	3	3	3	4	4	14	4	4	6	4	2	12	4	4	4	2	20	4.45
14	F	9	12	12	16	14	13	6	15	12	3	8	11	6	11	10	18	12	6	8	11	20	10.65
16	F	—	—	7	5	—	—	9	9	—	5	—	25	—	—	—	—	—	—	6	—	7	9.43
19	F	—	12	—	5	—	—	—	—	6	5	—	—	—	—	—	—	20	—	—	—	5	9.60
20	F	10	—	—	28	7	5	21	21	8	—	—	—	19	19	4	—	—	14	15	30	13	15.46
21	F	—	17	5	16	19	25	6	17	—	25	8	12	26	17	5	8	7	13	5	13	18	13.56
22	M	10	5	16	10	4	4	3	7	5	2	—	5	2	6	3	3	10	—	3	2	18	5.56
23	F	—	6	15	—	3	20	6	18	9	3	4	7	6	11	7	10	5	—	24	3	17	9.24
25	F	1	1	1	2	1	—	1	1	1	1	1	1	1	1	2	1	1	1	1	1	19	1.11
26	F	18	6	9	3	6	7	6	3	12	6	3	4	12	4	13	7	5	6	4	16	20	7.50
Frequency		7	8	11	11	11	11	13	12	12	12	10	12	12	10	13	11	11	10	11	12		
Mean Latency		8.4	7.6	10.7	9.5	8.6	10.5	8.6	10.4	8.3	8.3	5.4	8.9	10.4	10.4	8.2	10.9	9.5	9.6	8.8	11.1		

TABLE 2

DURATION IN SECONDS OF HALLUCINATORY RESPONSES ON 20 SUCCESSIVE EXTINCTION TRIALS *

The absence of response is indicated by a dash.

Subject No.	Sex	Trial No.																				Total Frequency	Mean Latency
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20		
1	M	—	—	3	3	2	2	1	3	3	2	7	4	5	3	3	5	3	4	4	3	18	3.33
3	M	—	—	—	—	2	1	2	—	—	—	3	—	—	—	2	1	—	—	—	0	7	1.57
6	F	—	—	—	—	—	—	—	—	2	1	—	1	3	1	2	3	—	1	—	2	9	1.78
10	F	—	—	—	—	—	4	—	—	—	—	—	—	—	—	1	3	1	2	—	—	5	2.20
11	M	—	—	7	5	3	—	2	1	1	1	1	1	1	—	1	—	—	—	1	—	12	2.08
12	F	2	—	3	—	—	4	1	2	2	—	3	4	3	—	—	—	1	4	—	1	12	2.50
13	F	6	8	4	14	19	8	8	7	12	1	0	1	3	4	18	0	1	25	33	1	20	6.54
14	F	5	5	4	7	0	4	12	9	4	3	5	2	2	5	4	4	2	6	2	1	20	4.30
16	F	—	—	1	5	—	—	1	1	—	2	—	5	—	—	—	—	—	—	1	—	7	2.29
19	F	—	1	—	3	—	—	—	—	4	4	—	—	—	—	—	—	5	—	—	—	5	3.40
20	F	3	—	—	0	2	10	1	0	1	—	—	—	0	1	2	—	—	1	2	1	13	1.85
21	F	—	3	2	3	2	2	2	1	—	3	2	3	2	3	2	3	1	2	3	4	18	2.39
22	M	2	3	2	8	4	4	5	3	3	3	—	3	3	4	2	3	1	—	4	4	18	3.39
23	F	—	3	3	—	5	7	2	3	3	3	4	2	2	2	5	4	2	—	3	0	17	3.12
25	F	1	1	1	1	1	—	1	0	1	2	2	1	1	4	2	1	1	1	1	0	19	1.21
26	F	4	4	4	2	2	1	7	3	1	3	3	3	2	2	3	3	4	8	4	5	20	3.40
Frequency		7	8	11	11	11	11	13	12	12	12	10	12	12	10	13	11	11	10	11	12		
Mean Duration		3.3	3.5	3.1	4.6	3.8	4.3	3.5	2.8	3.1	2.3	3.0	2.5	2.3	2.9	3.6	2.7	3.5	5.4	5.3	1.8		

* Duration is presented in gross units of 1 second. The units, or 'class intervals,' extend from the figure given up to but not including the next integer, *i.e.*, a duration presented as 0 extends to 0.99 seconds, and one presented as 1 extends from 1.00 to 1.99.

represents the differences in latency, duration, and frequency of response before and after the subjects were told that the tone would not be presented. All three of these differences are in a direction indicating that the tendency for the response to occur was weakened by the new instructions, and all of the differences are statistically reliable. A fourth measure of the strength of the response tendency, the proportion of subjects giving one or more hallucinatory responses

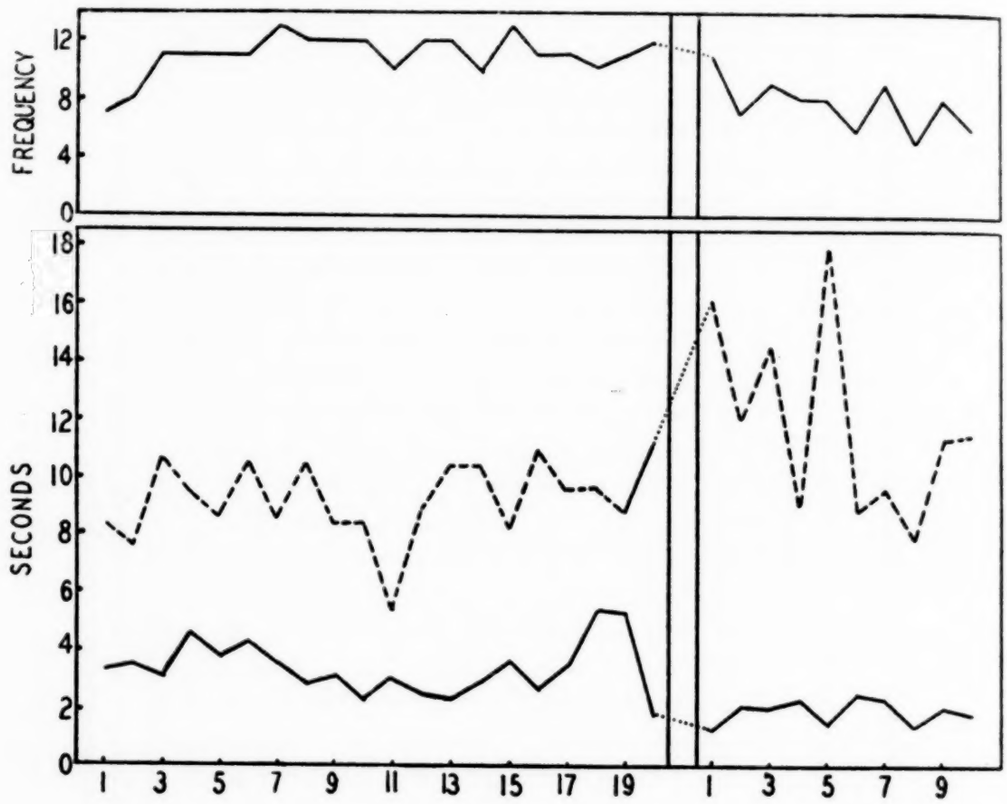

FIG. 1. The upper graph presents total frequency of hallucinatory responses on successive test trials. In the lower graph, the solid line indicates the mean duration, and the broken line indicates the mean latency of response. The two vertical lines (dotted lines in each curve) indicate the point at which the subjects were instructed that the tone would not be presented on following trials.

in ten test trials, is also presented in Table 4. Sixteen subjects gave one or more responses in the first and second ten extinction trials, and fifteen of the sixteen gave one or more responses in the ten test trials given after they were told not to expect the tone. This difference, while in the same direction with regard to strength of response tendency as the others, is not reliable.

The data from Table 3 are presented graphically in the right-hand portion of Fig. 1. There is some indication of a progressive trend in the curves for latency and frequency, and practically none for dura-

TABLE 3

LATENCY AND DURATION (IN SECONDS) OF HALLUCINATORY RESPONSES IN TEN 30-SECOND TEST TRIALS FOLLOWING INSTRUCTIONS THAT THE TONE WOULD BE OMITTED

The absence of response is indicated by a dash.

Latency

Subject No.	Sex	Trial No. 1	2	3	4	5	6	7	8	9	10	Total Frequency	Mean Duration	Answer on Questionnaire
1	M	9	—	5	7	—	8	5	2	8	6	8	6.25	—
3	M	—	—	—	—	—	—	—	—	—	—	0	—	—
6	F	17	—	10	—	9	8	6	28	19	—	7	13.86	—
10	F	—	25	22	—	—	—	—	—	—	—	2	23.50	A
11	M	—	16	—	11	20	—	—	—	4	—	4	12.75	C
12	F	23	—	—	17	—	—	10	—	—	4	4	13.50	A
13	F	26	—	10	10	—	7	28	—	15	—	6	16.00	B
14	F	4	4	8	4	8	2	5	5	4	13	10	5.70	A
16	F	16	—	—	—	30	—	4	—	—	—	3	16.67	C
19	F	11	—	—	—	—	—	—	—	—	—	1	11.00	A
20	F	—	—	—	—	—	—	16	—	—	—	1	16.00	C
21	F	8	21	28	17	13	18	—	—	19	17	8	17.63	A
22	M	25	—	—	—	—	—	—	—	—	8	2	16.50	A
23	F	—	4	28	5	29	—	4	—	—	21	6	15.17	—
25	F	11	2	11	1	29	—	8	1	18	—	8	10.13	B
26	F	27	12	8	—	5	10	—	3	3	—	7	9.71	C

Duration

Subject No.	Sex	Trial No. 1	2	3	4	5	6	7	8	9	10	Total Frequency	Mean Duration	Answer on Questionnaire
1	M	1	—	2	1	—	4	2	2	2	1	8	1.88	—
3	M	—	—	—	—	—	—	—	—	—	—	0	—	—
6	F	2	—	3	—	1	2	2	1	2	—	7	1.86	—
10	F	—	2	1	—	—	—	—	—	—	—	2	1.50	A
11	M	—	2	—	3	2	—	—	—	3	—	4	2.50	C
12	F	1	—	—	2	—	—	1	—	—	1	4	1.25	A
13	F	1	—	3	5	—	3	10	—	4	—	6	4.33	B
14	F	3	4	4	3	2	3	4	1	2	3	10	2.90	A
16	F	0	—	—	—	0	—	0	—	—	—	3	0.00	C
19	F	1	—	—	—	—	—	—	—	—	—	1	1.00	A
20	F	—	—	—	—	—	—	2	—	—	—	1	2.00	C
21	F	2	1	2	1	3	2	—	—	1	2	8	1.75	A
22	M	1	—	—	—	—	—	—	—	—	2	2	1.50	A
23	F	—	1	1	2	0	—	0	—	—	2	6	1.00	—
25	F	1	2	2	1	1	—	1	2	1	—	8	1.39	B
26	F	1	3	0	—	3	1	—	1	1	—	7	1.43	C
Total Freq.		11	7	9	8	8	6	9	5	8	6			
Mean Lat.		16.1	12.0	14.4	9.0	17.9	8.8	9.6	7.8	11.3	11.5			
Mean Dur.		1.3	2.1	2.0	2.3	1.5	2.5	2.4	1.4	2.0	1.8			

tion. The reliabilities of these trends were tested by obtaining critical ratios for the differences between the means for the first and second five trials (averages of subject means in each case). The

TABLE 4

MEAN FREQUENCY, LATENCY, DURATION, AND PERCENTAGE OF SUBJECTS GIVING ONE OR MORE RESPONSES (IN 10 TRIALS) IN THE 20 TRIALS PRECEDING AND THE 10 TRIALS FOLLOWING THE INSTRUCTION OF THE SUBJECTS NOT TO EXPECT THE TONE

The mean frequency for the 20 preceding trials has been divided by 2 in order to make it comparable to that for the 10 following trials.

	Mean Frequency (per 10 trials)	Mean Latency	Mean Duration	Percentage of Subjects Responding
Trials 1–20 (preceding)	6.88	10.19	2.83	53%
Trials 1–10 (following)	4.81	13.62	1.75	50%
Difference	2.07	3.43	1.08	3%
Diff./$\sigma_{Diff.}$	5.49	3.57	4.92	0.23

critical ratios obtained were as follows:

Critical Ratio for Latency = 2.65
Critical Ratio for Frequency = 1.92
Critical Ratio for Duration = 1.08

Only the differences for latency and frequency approach statistical significance. The meaning of the trends is also unclear, since the two stronger trends were both decreasing. A decrease in latency is ordinarily interpreted as an *increase* in the response tendency, while a decrease in frequency is interpreted as a *decrease*. There is no obvious hypothesis which will explain this apparent contradiction.

DISCUSSION

The Failure to Find Evidence of Extinction.—There are reasons both for expecting and for not expecting extinction to occur when the tone is omitted in the present experiment. The omission of the UCS has resulted in the gradual disappearance of the response (extinction) in the majority of conditioning experiments in which the response was established to a neutral stimulus by a pairing procedure. Included in this majority are the results obtained by Bogoslovski (**2**) in the only sensory conditioning experiment in which extinction was studied.

On the other hand, a number of experiments have failed to demonstrate the occurrence of extinction (**1**, **7**, **8**, **11**, **14**, **15**, **17**). A fact which may be significant is that all but two of these (7 and 8, which used monkeys as subjects) were performed with human subjects. In a summary of such experiments, Hilgard and Marquis suggest that ". . . an expectancy may be established by self-instructions and this takes the place of the original reinforcement which is now omitted" (**9**, p. 122). This interpretation is supported by the results of an animal study by Tolman and Brunswik (**18**) in which it

was found that rats responded to a stimulus situation which sometimes led to shock [*i.e.*, in which they 'expected' shock] in the same way that they responded to one which always led to shock. It would be assumed, since expectation is presumably an important factor in the present experiment, that expectation of the tone would substitute as a reinforcement for the actual presentation of the tone. This argument is weakened by the fact that Humphreys (**13**) found than an expectation in human subjects is itself subject to 'extinction' when the expected stimulus is repeatedly and consistently omitted.

It should be pointed out, however, that in the present experiment it may be incorrect to state that the UCS was omitted. If the condition of reinforcement is assumed to be 'hearing the tone' rather than "hearing the tone which is produced externally," then for those subjects who gave one or more hallucinatory responses in the 'extinction' series, the reinforcement was not consistently omitted, and for three subjects who gave twenty responses in the twenty extinction trials, it was never omitted. Since, under the conditions of the present experiment, the hallucinated tone is not distinguished by the subjects from the perception of a physical tone, it would seem quite possible that this should be the case.[3] Results obtained by Humphreys (**13**) indicate that the fact that the frequency of reinforcement in the present experiment is reduced from 100 percent in the training series to less than 100 percent in the 'extinction' series for all but three of the subjects does not contradict this hypothesis. In the experiment mentioned above, Humphreys showed that reinforcement on only 50 percent of trials during training was as effective, if not more so, than 100 percent reinforcement in conditioning a verbal expectation, and he has obtained similar results for partial reinforcement in an eyelid conditioning experiment (**12**). In a study of a similar problem using dogs, in which a flexion response was conditioned by the use of shock, Brogden (**3**) found no difference in CR frequency until the percentage of reinforcement was reduced to 20 percent. Consequently, the failure of extinction to occur may be explained by assuming that the hallucination of the tone, once acquired as a result of the training or conditioning procedure, is as effective in reinforcement as the perception of the actual physical tone. As long as the training is sufficient to assure an initial proportion of hallucinatory responses in the extinction series above some minimum, then the hallucination itself serves as a reinforcement, and extinction should not be expected.

The Effect of Instructing Subjects Not to Expect the Tone.—Three

[3] The fact that extinction does not occur when the UCS is hallucinated has important implications for the problem of the nature of reinforcement. It suggests that a response rather than a stimulus is the necessary condition for reinforcement in conditioning.

measures, frequency, latency, and duration, which are ordinarily used as indices of the strength of a response tendency, gave reliable differences indicating that instructing subjects not to expect the tone weakened the tendency for the hallucination to occur in response to the light. A fourth measure, the proportion of subjects giving one or more hallucinatory responses in ten test trials, presented a small and unreliable difference in the same direction. These results are similar to those found in the sensory conditioning experiment by Bogoslovski (**2**) who conditioned a change in electrical sensitivity of the dark-adapted eye. He found that the CR almost completely disappeared on trials in which he told subjects that the light would not be presented. The CR was present at full strength on the day preceding and the day following that in which the no-light instructions were given, although no intermediate training was given to reinstate it.

Similar results have been obtained in other conditioning experiments by Cole (**4**), using the eyelid response, and by Cook and Harris (**5**) and Mowrer (**16**) using the GSR. In Cole's eyelid experiment, extinction was greatly facilitated by instructing the subjects that the UCS would be omitted. The GSR almost completely disappeared in the other studies when the subjects were told that the shock would not be presented. These results are supplemented by the work of Hilgard, R. K. Campbell, and Sears (**10**), who found that conditioned discriminations were more rapid when the subject received instructions concerning the positive and negative stimuli. We may conclude that the hallucinatory response studied in the present experiment resembles a conditioned response in being weakened when the subject is instructed not to expect the UCS.

However, in many of these studies, including the present one, there is evidence that although extinction is facilitated, it is not immediately complete when the subject is instructed not to expect the UCS. In the present experiment, fifteen of the sixteen subjects (50 percent of the total group of thirty S's) told not to expect the tone gave one or more hallucinatory responses in ten test trials after these instructions were given. This may be compared with the 20 percent of sixty subjects who gave one or more responses in ten similar test trials given prior to any training in the previous experiment (6, Control Group III, Table 6). There are two interpretations of the failure of occurrence of immediate, complete extinction when subjects are told not to expect the UCS. The conditioning process may produce a residue of response tendency which is not dependent on the subject's verbal expectation, or it is possible that good rapport is not attained, so that the subject continues to expect the stimulus in spite of instructions that it will not occur. Some evidence on this

point was obtained in connection with the present experiment. Several weeks after the last subject had been run, a rumor was encountered indicating that some of the subjects had not believed the experimenter when he told them that the tone would not be presented on the subsequent trials.[4] The questionnaire presented below was sent to those of the sixteen subjects giving responses who were still available, and twelve were returned. The number of subjects checking each alternative in the questionnaire is indicated in the space supplied for checking.

Questionnaire given to subjects who gave hallucinatory responses:

> Those taking part in the auditory threshold experiment will remember that toward the end of the hour a few trials were given in which you were told that no tone would be presented. There has been some indication that some of the subjects believed that this was some sort of trick and that the tone actually was presented. Since it is important for the interpretation of the results to know whether the subjects did or did not expect a tone during these trials, will you please indicate whether you believed the experimenter when he told you that the tone would not be presented. Please indicate with a check those statements which describe your attitude *at the time of the experiment.*
>
> *A.* Believed experimenter—did not expect tone..... (6)
> *B.* Believed experimenter at the time but later thought the tones you heard were not imagined..... (2)
> *C.* Did not believe experimenter—expected the tone to be presented as before.... (4)
> *D.* Do not remember..... (0)
> In case you are interested, the tone was *not* presented during these trials.

The answer checked on the questionnaire by each of the twelve subjects is indicated by an 'A,' 'B,' or 'C' at the right in Table 3. The number of cases involved is too small for statistically reliable differences to occur, but careful inspection suggests that the subjects' expectations affected the response. When frequency per ten test trials, mean latency, and mean duration for each subject are compared for the trials before and after instructions not to expect the tone, there are five cases in which the difference for the individual subject is in a direction opposite to that for the group (*i.e.*, the response tendency was stronger after the instructions not to expect the tone). Only one of these discrepancies is found for a subject reporting that he believed the experimenter's instructions. Mean latency for Subject No. 14, who gave an 'A' answer on the questionnaire, is shorter for the last ten than for the first twenty test trials. The remaining four discrepancies are given by the six subjects giving 'B' or 'C' answers indicating more or less disbelief. Subject No. 25 ('B' answer) gave a mean duration which was longer for the last ten, Subject 11 ('C') gave a shorter mean latency and a longer mean duration,

[4] This gives some indication of the 'reality' of the hallucination to the subjects. The instructions that the tone would not be presented were given immediately after a series of twenty trials in which the tone had been entirely omitted. In spite of this and the fact that no tone was presented in the next ten trials, six subjects indicated that they did not believe the new instructions.

and Subject 20 ('C') gave a longer mean duration in the last ten tests. The distribution of these individual-vs.-group discrepancies suggests that the hallucinatory response tendency is weakened to a greater extent by instructions not to expect the tone when the subject believes these instructions than when he does not.

Summary

Thirty subjects were given training by means of a procedure developed in a previous experiment in order to produce an auditory hallucination of a 1000-cycle tone. The sixteen subjects who gave hallucinatory responses were given twenty test trials in which the cue stimulus (light) was presented, but the tone was omitted. No indication of experimental extinction was found in either frequency, latency, or duration of response. Instructing the subjects that the tone would be omitted resulted in reliably reduced frequency and duration, and increased latency in the following ten test trials. Results of a questionnaire suggested that the instructions were less effective for those subjects who did not believe them and continued to expect the tone. None of the evidence obtained indicates functional relationships different from those to be expected on the basis of results of previous conditioning experiments.

(Manuscript received October 11, 1940)

References

1. Baker, L. E., The pupillary response conditioned to sub-liminal auditory stimuli, *Psychol. Monogr.*, 1938, **50**, No. 223, 32 pp.
2. Bogoslovski, A. I., An attempt at creating sensory conditioned reflexes in humans, *J. exp. Psychol.*, 1937, **21**, 403–422.
3. Brogden, W. J., The effect of frequency of reinforcement upon the level of conditioning, *J. exp. Psychol.*, 1939, **24**, 419–431.
4. Cole, L. E., A comparison of the factors of practice and knowledge of experimental procedure in conditioning the eyelid response of human subjects, *J. gen. Psychol.*, 1939, **20**, 349–373.
5. Cook, S. W., & Harris, R. E., The verbal conditioning of the galvanic skin reflex, *J. exp. Psychol.*, 1937, **21**, 202–210.
6. Ellson, D. G., Hallucinations produced by sensory conditioning, *J. exp. Psychol.*, 1941, **28**, 1–20.
7. Harlow, H. F., Experimental analysis of the role of the original stimulus in conditioned responses in monkeys, *Psychol. Rec.*, 1937, **1**, 62–68.
8. Hilgard, E. R., & Marquis, D. G., Conditioned eyelid responses in monkeys, with a comparison of dog, monkey, and man, *Psychol. Monogr.*, 1936, **47**, No. 212, 186–196.
9. Hilgard, E. R., & Marquis, D. G., *Conditioning and learning*. New York: Appleton-Century, 1940.
10. Hilgard, E. R., Campbell, R. K., & Sears, W. N., Conditioned discrimination: the effect of knowledge of stimulus-relationships, *Amer. J. Psychol.*, 1938, **51**, 498–506.
11. Hudgins, C. V., Conditioning and voluntary control of the pupillary light reflex, *J. gen. Psychol.*, 1933, **8**, 3–51.
12. Humphreys, L. G., The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions, *J. exp. Psychol.*, 1939, **25**, 141–158.

13. Humphreys, L. G., Acquisition and extinction of verbal expectations in a situation analogous to conditioning, *J. exp. Psychol.*, 1939, **25**, 294–301.
14. Marquis, D. G., & Porter, J. M., Jr., Differential characteristics of conditioned eyelid responses established by reflex and voluntary reinforcement, *J. exp. Psychol.*, 1939, **24**, 347–365.
15. Menzies, R., Conditioned vasomotor responses in human subjects, *J. Psychol.*, 1937, **4**, 75–120.
16. Mowrer, O. H., Preparatory set (expectancy)—a determinant in motivation and learning, *Psychol. Rev.*, 1938, **45**, 62–91.
17. Razran, G. H. S., Conditioned responses: an experimental study and a theoretical analysis, *Arch. Psychol.*, N. Y., 1935, **28**, No. 191, pp. 124.
18. Tolman, E. C., & Brunswik, E., The organism and the causal texture of the environment, *Psychol. Rev.*, 1935, **42**, 43–77.

A NOTE ON THE RELATIONSHIP BETWEEN 'PERSONALITY' AND THE ALPHA RHYTHM OF THE ELECTROENCEPHALOGRAM[1]

BY CHARLES E. HENRY
Bradley Home and Brown University
AND
JOHN R. KNOTT
Departments of Psychology and Neurology
University of Iowa

It has been reported by several investigators (Gottlober, 2; Lemere, 4; Saul, Davis and Davis, 5) that there is a relationship between what is loosely called 'personality type' and one rather limited aspect of the electroencephalogram: the percentage of time an alpha rhythm is present in the EEG record.

Although none of these investigations has been highly quantitative, one has attempted to relate qualitative estimates of personality and quantitative determinations of percent time alpha. Gottlober studied the EEGs of 67 'normal' individuals who were selected by subjective rating to be either 'extremely extroverted' or 'extremely introverted.' (In addition the subjects filled out the Nebraska Personality Inventory (3) and Gottlober reported 75 percent agreement between subjective ratings and ratings derived from the inventory scores.) He concluded from his investigation that "an organismic pattern in which a dominant-subdominant rhythm [2] appears accompanied by an extroverted personality, as determined by our methods of rating, will be more common than one in which it appears accompanied by an introverted personality."

The qualification, 'by our methods of rating,' renders the most quantitative of these previous studies essentially unrepeatable and therefore makes objective, verifiable data pertaining to the problem the next essential step in the continuation of such research. The present study was undertaken to provide such data.

Apparatus and Procedure.—The apparatus consisted of a high-gain, three-stage, resistance-capacitance-coupled amplifier which drove an ink-writing oscillograph. The leads to the subject were flattened lead disks, one secured to the left occipital region, the other to the lobe of the left

[1] The present report is one of a series of papers under the direction of the junior author, dealing with physiological indexes of psychological phenomena. All of the data were collected and evaluated in the laboratory of physiological psychology, in the Department of Psychology, University of Iowa.

[2] This refers to a percent time alpha over 50. When less than 50, he labelled the rhythm 'mixed-rare.'

ear. The subjects lay, eyes closed, in an electrostatically shielded, dark, soundproofed room. A minimum of five minutes of continuous record was taken from each subject; five minutes total record was read for each subject.

It should be pointed out that these conditions vary in two respects from Gottlober's. First, we recorded with an ink-writing oscillograph, while he recorded with a Westinghouse optical oscillograph. We have no reason to believe that this variation could influence the data to be considered, since the frequency response of the writer was linear throughout the band of frequencies studied. Second, our records were *five minutes* in duration, while his were only *thirty seconds* in duration. Due to the transient characteristics of the alpha rhythm in many subjects, the longer intervals appear to provide the most stable measures of the rhythm.

The method of computing the alpha index used here was to mark off all sections of the record on which waves with a frequency of eight to twelve cycles per second and greater than seven microvolts in amplitude appeared. This total was then expressed as the percentage of the entire record.

Subjects.—Eighty individuals, forty-one males and thirty-nine females, who were attending undergraduate and graduate courses in speech and psychology, and who volunteered to serve as subjects, were studied. The source of subjects was essentially the same as the source of Gottlober's subjects. Since Gottlober had administered the Nebraska Personality Inventory to his subjects, these eighty persons were given the same inventory to fill out. On the basis of the norms supplied with it at the time this investigation was undertaken, we labelled all those individuals below Q2 'extroverts,' and all those above that point 'introverts.' The alpha indices were determined, and all those subjects above 50 (percent) were labelled 'high alphas' and all those below that point, 'low alphas.'

Results

Of the 80 subjects studied, 55, or 69 percent, were extroverted, and 24, or 44 percent, were high alphas. Forty-three percent of the extroverts were high alphas, and 60 percent of the introverts were high alphas.

This tendency reverses the tendency reported by Gottlober. To discover whether the apparent relation between introversion and a high alpha index is statistically significant, the ratio of the difference to the standard error of the difference was computed and found to be 1.1. This does not indicate the difference to be significant.

The x^2 test yielded a P of .15. Thus, we would not be justified in rejecting the hypothesis that the differences in the distributions are due entirely to chance fluctuations. In other words, in this set of data, no relationship between personality, as measured, and the percent time alpha of the EEG is indicated.

Discussion

In an attempt to evaluate our negative data in terms of the apparently positive data of Gottlober, the latter are presented in some detail.

Gottlober's conclusion that "an organismic pattern in which a dominant-subdominant rhythm appears accompanied by an extroverted personality as determined by our methods of rating will be more common than one in which it appears accompanied by an intro-

verted personality" was based upon the following data:

82 percent of the judged extroverts were high alphas
18 percent of the judged extroverts were low alphas
57 percent of the judged introverts were high alphas
43 percent of the judged introverts were low alphas

The difference between 82 and 18 percent proved to be statistically significant. Hence Gottlober concluded that 'extroversion' went hand in hand with a high alpha index. The difference between 57 and 43 was not.

When data based upon personality groups upon which both the judges' ratings and the inventory scores agreed were similarly analyzed, the same trends were substantiated.

When we analyzed Gottlober's inventory data it was found, using this original method of comparison, that a significantly greater percentage of his self-rated extroverts exhibited a high alpha index than a low alpha index. These percentages were 77 and 23, respectively, and the critical ratio was 5.7. Sixty-four percent of the self-rated introverts were high alphas, as opposed to 36 percent who were low alphas. The critical ratio was 2.2.

These comparisons, however, are neither statistically legitimate nor pertinent to the conclusion which he drew from them. What must be discovered is whether or not the proportion of extroverts exhibiting high alpha indexes is significantly greater than the proportion of introverts exhibiting high alpha indexes.

Gottlober presented these data but did not discuss them. He reported the difference between the percentage of extroverts exhibiting high alphas and the percentage of introverts exhibiting high alphas (as extroversion or introversion was rated by his judges) to be 25, and the ratio of this to its standard error he reported to be 2.3. This indicates a marginal degree of significance which is suggestive rather than final.

When subjects were selected for comparison on whom the judges' ratings and the inventory ratings were in agreement, he reported the difference to be reduced to 18 percent, and the ratio was 1.5. This latter ratio is not considered significant. (In applying the x^2 test to these data we had recourse to original data and found the difference to be 25 percent, and the ratio to be 1.9, which is not significant.)

We analyzed his inventory data (independent of judges' ratings) and found the difference to be 13 percent, and the ratio of the difference to its standard error to be 1.1, which does not indicate the difference to be significant.

Application of the x^2 test to these three sets of data yielded Ps equal to .03, .05, and .25, respectively. In the first two sets of data

there is a relationship between extroversion and a high alpha index indicated at the 3 and 5 percent levels of confidence (that is, the hypothesis that the relationship thus indicated is attributable to chance may be rejected at those levels of confidence).

In order to measure the degree or magnitude of the relationships which were indicated to be significant by x^2, bi-serial correlations were computed between the variables. Using the judges' ratings only, $r_{bis} = -.38 \pm .06$. Using data obtained from the agreement of judges' ratings and self-ratings, $r_{bis} = -.36 \pm .11$.

It appears, therefore, in the original data of Gottlober, that the more objective and operational the definition of personality became, the less likely was there need to reject the hypothesis that the relationship between an electroencephalographic variable (in this case, the percent time alpha) and personality was a chance relationship. To make a further test of this, the inventory data of Gottlober were combined with the inventory data obtained in the present investigation. When this was done, it was found that of the 147 subjects:

a. 58 percent were high alphas; of these, 62 percent were extroverted.

b. 42 percent were low alphas; of these, 67 percent were extroverted.

c. 64 percent were extroverted; of these, 57 percent were high alphas.

d. 36 percent were introverted; of these, 62 percent were high alphas.

It appears, therefore, that there was an almost equal proportion of extroverts in *both* the high and the low alpha groups; and that there was an almost equal proportion of high alpha subjects in *both* the extroverted and the introverted groups. None of the pertinent comparisons indicated the differences to be statistically significant.

However, the obvious fact demanding discussion is that the x^2 test applied to Gottlober's judges'-rating data indicates a probably significant correspondence to exist between a high alpha index and extroversion. We are necessarily required to ask, however, *what* criteria of the judges were met to warrant the labelling of a particular individual as either extroverted or introverted. The answer to this question, unfortunately, is not available.

The inventory (self-rating) data of Gottlober do not support his contention that "an organismic pattern in which a dominant-subdominant rhythm appears accompanied by an extroverted personality, as determined by our methods of rating, will be more common than one in which it appears accompanied by an introverted personality."

The qualification, ' by our methods of rating' (*i.e.*, judgments),

makes it virtually impossible to justly challenge the statement, since the same methods were not used in the present investigation. We can ask, however, that the methods be at least qualitatively described.

Our own sample, categorically divided into extroverts and introverts on the basis of the self-ratings, do not support the conclusions drawn from Gottlober's data, either judges'-rating of self-rating. This is to be stressed, since he himself reported 75 percent agreement between test and judges.[3] As a matter of fact, any indications which our sample might make would be opposite to the direction of his sample, for there was a trend (insignificant, of course) for there to be fewer high alpha extroverts and more high alpha introverts.

The combined self-rating data seem to lead us to the conclusion that there is no relationship between 'personality' (thus measured) and the magnitude of the alpha index. The percentage of high alphas in the two personality groups is very nearly equal to the percentage of high alphas (64 percent) in an unselected population of 401 normal subjects, reported by H. Davis (1).

It might appear, therefore, since Gottlober's sample was loaded with high alphas (72 percent of all subjects), and since both samples had more extroverts than introverts, that the 'relationship' is in the last analysis attributable to chance factors alone.

Until further (and, we would like to add, objective, operational) data are forthcoming, we suggest that this possibility be regarded as a qualification of the original conclusions of Gottlober. We would further suggest, in view of the relationship found possibly to exist between the judges' ratings of personality and the alpha index, that additional research start at that point and attempt to uncover the specific (*i.e.*, objective and operational) behavioral items, or groups of items, associated with a given personality rating and a given electroencephalographic pattern.

(Manuscript received October 8, 1940)

References

1. Davis, H. The electroencephalogram. *Tabulae Biologicae*, 1938, **16**, 116–131.
2. Gottlober, A. B. The relationship between brain potentials and personality. *J. exp. Psychol.*, 1938, **22**, 67–74.
3. Guilford, J. P., & Guilford, R. B. Personality factors *S*, *E* and *M*, and their measurements. *J. Psychol.*, 1936, **2**, 109–127.
4. Lemere, F. The significance of individual differences in the Berger rhythm. *Brain*, 1936, **59**, 366–375.
5. Saul, L. J., Davis, H., & Davis, P. A. Correlations between electroencephalograms and the psychological organization of the individual. *Trans. Amer. Neurol. Assoc.*, 1937, **63**, 167–169.

[3] We subjected his comparative data to correlation with the bi-serial *r*, and found the coefficient to be .62 ± .02. Hence we feel that though agreement between tests and judges was not perfect, it was sufficiently high to make a comparison of the two 'definitions' of personality possible.

A NOTE ON CHRONAXIC TECHNIQUE

BY JACQUES S. GOTTLIEB AND O. D. FOWLER

Iowa State Psychopathic Hospital

There have been a number of investigations of chronaxie in the clinical field. Most of these have attempted to study the relationships between chronaxies with electrodes as described by Bourguignon (1) in his classical work on the chronaxie in pathological cases. Certainly there are well defined differences distinguishable by such methods, in those cases where there is discernible gross pathology of the nervous system or of muscle. Thus, the chronaxie may be found to be lengthened enormously in cases of degeneration of the peripheral nerve, and the progress of the regenerative process may be noted by following the progressive shortening of this measure as time goes on.

Such studies, and the more delicate ones done upon physiological preparations with the theoretical implications thus raised, have led a number of experimenters to attempt clinical studies upon man, in cases where there is no gross pathology demonstrable. Workers in the field of functional disorders of the nervous system have contributed many such studies, and these have been supplemented by studies upon normal human beings; other workers have varied these approaches by the use of drugs. These experiments have led to conflicting results and in many cases have brought criticism both of the methods used and of the inferences drawn.

The present paper will attempt to discuss some of the points to be considered in determining whether or not such methods are valid, with respect to some findings of the present authors in their attempts to apply chronaxie measures to clinical cases where there is no demonstrable organic pathology.

Much controversy has arisen concerning the nature, size, and manner of disposition of the stimulating electrodes, both in studies upon man and in experimentation upon physiological preparations. Many of the facts revealed have received too little attention in clinical studies. For example, the electrodes described by Bourguignon have been widely used, and discussion of results, or lack of results, when using these methods have concerned only theoretical implications of chronaxies, with little consideration of the fact that only gross differences, such as found in organic lesion cases, may be detected by such means.

The active electrode referred to is a chlorided silver button covered with canton flannel and gauze and soaked in physiological saline solution. In use, the electrode rests upon the moistened surface of the

skin over the motor point of the muscle, usually being held there by means of an attached handle in the operator's hand. Thus there are the following sources of error:

1. *Variations in pressure*, even slight ones due to the tremor of the operator's hand, affect the magnitude of the measures taken. Distortion of the surfaces, variations in the effective electrode area, and changes in the amount of interposed inactive tissue are important factors here.

2. *Shifts in position* of the electrode upon the skin, however slight, may serve also to bring about changes due to similar factors.

3. *Moisture changes* of the electrode and of the skin surface, due to evaporation or spread of the saline solution, result in alteration of the shape and size of the effective electrode, thus causing rapid variation of the measures taken, an important source of error frequently overlooked. Many studies have been carried through in which the electrodes were moistened from time to time, a necessary procedure to maintain the effective contact but one which serves only to point out the existence of this source of error.

4. *The indication of response* taken as a threshold reaction is frequently open to serious objections, as in the common procedure of observing the visible contraction of the muscle. Apart from the fact that the reaction will vary with the uncontrolled electrode factors, such a method of threshold determination is obviously very crude and subject to changes with light, and slight movements of position. In many subjects difficulty is experienced because of the disposition of the subcutaneous structures, especially layers of fat.

5. *The exact application* of the stimulus provides a difficult problem which is not easily surmounted experimentally. The existence of interposed tissue makes stimulation through the skin indefinite as regards the number and location of the nerve or muscle fibers stimulated. Such stimulation is always of a complex nature, involving a number of motor units. Since periodic tonic discharges are normally occurring, changes in refractory phase are also taking place, virtually assuring that otherwise controlled electrical stimulation cannot be repeated under exactly the same conditions except by chance. There is also the added possibility of sensory stimulation with reflex effects influencing the status of the fibers in successive periods.

Procedure

With the object of eliminating as many of the known uncontrolled factors mentioned as possible, a new electrode was designed and tactile observation of the threshold response was used.

The electrode used is a rubber cap with a flange base, having a silver wire sealed through its side and connected to an insulated

flexible copper wire lead. That portion of the wire inside the nipple is in spiral form and is electrolytically chlorided before use. There is a small hole (2 mm or less) in the top of the cap.

In use, the motor point of the muscle is first found by means of an ordinary small hand electrode of the Bourguignon type, having the same electrode diameter as the inner diameter of the rubber cap. After the point has been carefully marked and the surface of the skin throughly dried, the base of the cap is moistened with collodion and carefully placed over the marked point where it is held until securely dried. The cap is then filled with physiological saline solution through the aperture in the top of the cap by means of a hypodermic syringe. It thus provides a moist contact surface of uniform size and pressure, which needs no further attention during the experimental period.

As an indication of response, the contraction of the muscle as determined by the sense of touch of the operator's fingers was used. The electrode was applied over the superior motor point of the radial nerve of the human forearm, with the subject reclining in a relaxed position.

Results

1. The consistency of the threshold magnitude during successive determinations within a few minutes' time is much greater with the type of electrode described than with the classical Bourguignon type. The following comparable sets of determinations illustrate this point:

Intensity Thresholds—Obtained with Electrode Described (volts)	Intensity Thresholds—Obtained with Bourguignon Type Electrode (volts)
71	75
69	90
70	63
72	71
71	74
71	82

2. The touch method of threshold determination is superior to the visual method frequently employed. After the threshold value is obtained, a variation of one or two volts will practically always serve to abolish or reëstablish the response as determined by this method; time factor measurements show like consistency. With the visual method there is marked variation in consistency of determinations with different subjects; in some cases it was found possible to change the 'threshold' value of intensity by as much as fifteen or twenty volts before the observer could detect the variation.

(Manuscript received October 7, 1940)

References

1. Bourguignon, G. *La chronaxie chez l'homme.* Paris: Libraires de l'Académie de Médecine, 1923.

A STIMULUS TIMER OPERATING WITHOUT COMMUTATOR CONTROL

BY STEVENSON SMITH

University of Washington

This timer is of value in that it permits the subject to get ready to receive a stimulus and, when ready, to administer the stimulus to himself without waiting. In the motor driven commutator timer the subject must wait for the beginning of the next time interval. This is a disadvantage in certain forms of visual stimulation where visual accommodation must be optimum when the stimulus is received. The

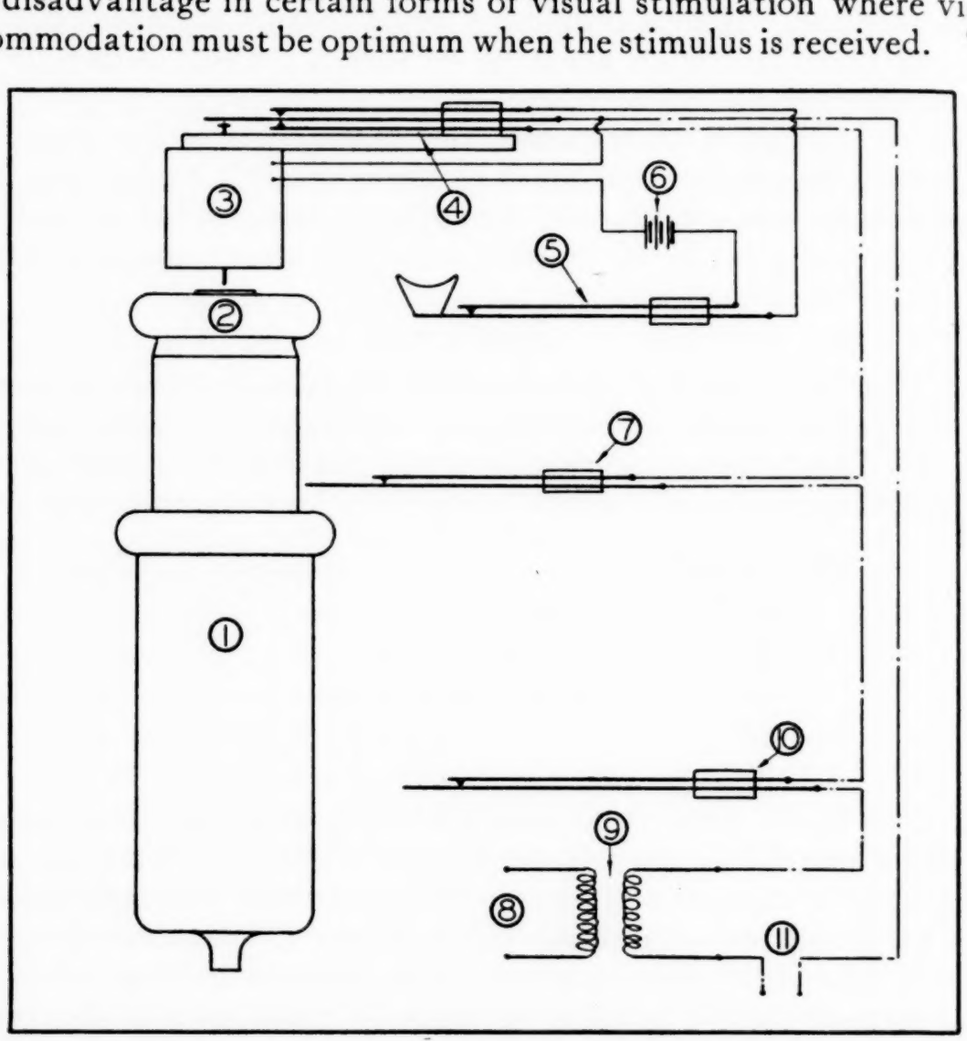

FIG. 1

timer here described is controllable by the subject in the same way that a fall-hammer or a camera shutter is controllable. It has the further advantage of being silent.

The timing element is a 50 cc Luer glass syringe with accurately ground contact surfaces manufactured by Becton, Dickinson and

Company, Rutherford, New Jersey. When held vertically the piston of the syringe falls of its own weight as the air escapes from the nozzle. Any distance within the range of piston's fall may be calibrated for a corresponding time interval. In operation the falling piston closes a circuit and then opens it. In the upper position the piston (2) is held by a small electro magnet (3) acting on a small iron washer cemented to the handle. Through the magnet is a small movable brass pin which works switch (4). In the upper position the piston holds switch (4) so the 110V. (11) circuit is open and the 6V.D.C. (6) circuit is closed. When the subject's key (5) is opened the magnet circuit is broken and the piston falls, breaking the D.C. contact and closing the 110V.D.C. circuit. When the piston arrives at another position in its fall it strikes and opens switch (7) which again opens the circuit. In order to keep the 110V. circuit open while the piston is being raised, switch (10) is opened by the same hand operated lever that sets the piston for the next operation. The D.C. circuit is arranged so that the magnet is energized only when the plunger is at the top. This is done by the upper leaf of switch (4). Switches (4), (5), (7), and (10) are made from spring bronze strips and mounted on bakelite spacers. The time interval is changed by raising and lowering magnet (3).

The constancy of the instrument is sufficient for most uses, as is indicated by the results of calibrating it on a 60 cycle time line recorder at the time interval at which it has been in use. The mean reading error was .0009 seconds. In sixty-two consecutive operations, the mean time interval being .0146 seconds, the S.D. of the distribution was .0017 seconds. For longer intervals the variability would be greater. The instrument was constructed by George Price, department shop technician.

(Manuscript received October 21, 1940)